

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:11 ; Search time 85 Seconds
(without alignments)
1092.412 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103

Sequence: 1 DAHKEVAHREFKDLGENFK.....TCFAEGKLVAAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq 19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	3103	100.0	585 10	AAAP90388 Mature human serum
2	3103	100.0	585 11	AAAR05318 Human serum albumi
3	3103	100.0	585 11	AAAR08457 Human serum albumi
4	3103	100.0	585 16	AAAR0301 Human serum albumi
5	3103	100.0	585 18	AAAO20111 HSA protein sequenc
6	3103	100.0	585 21	AAAY84873 Amino acid sequenc
7	3103	100.0	585 21	AAAY83946 yeast codon-biased
8	3103	100.0	585 22	AAAB79006 Human mature album
9	3103	100.0	585 22	AAAE13399 Human albumin (HA)

10	3103	100.0	585 22	AAAM52567 Mature human serum
11	3103	100.0	585 22	AAEL13129 Human albumin (HA)
12	3103	100.0	585 22	AAEL12403 Human albumin (HA)
13	3103	100.0	585 22	AAE08578 Human serum albumi
14	3103	100.0	585 23	ABG71291 Glycosylated prote
15	3103	100.0	585 23	ABG63321 Human serum albumi
16	3103	100.0	585 23	ABG700986 B lymphocyte stimu
17	3103	100.0	585 23	ABG33847 Human B lymphocyte
18	3103	100.0	585 23	AAU75220 Mature form of hum
19	3103	100.0	609 21	AAAB36542 Recombinant human
20	3103	100.0	609 21	AAAB36549 Recombinant human
21	3103	100.0	609 21	AAAY78147 Pre human serum al
22	3103	100.0	609 24	ABU57252 Human serum albumi
23	3103	100.0	609 24	ABU57253 Human serum albumi
24	3103	100.0	610 14	AAAR39510 Chimeric human ser
25	3103	100.0	616 24	AAE30916 Val8-GLP-1-human s
26	3103	100.0	624 24	AAE30919 Human serum albumi
27	3103	100.0	631 24	AAE30917 Val8-GLP-1-linker-
28	3103	100.0	640 24	AAE30918 Gly8-GLP-2-GLP-1-C
29	3103	100.0	640 24	AAE30920 Extensin-4-linker-h
30	3103	100.0	670 21	AAAB36543 Recombinant human
31	3103	100.0	670 21	AAAB36550 Prepro-HSA-G-CSF c
32	3103	100.0	783 14	AAAR39473 G-CSF-(Gly)4-HSA c
33	3103	100.0	787 14	AAAR39477 HSA-vWF(470-713) f
34	3103	100.0	853 14	AAAR39472 Sequence of mature
35	3099	99.9	585 10	AAAP93344 Mature protein of
36	3099	99.9	585 19	AAWS9841 Human serum albumi
37	3099	99.9	608 17	AAAR96229 Sequence of human
38	3099	99.9	609 4	AAAP30089 Human serum albumi
39	3099	99.9	609 17	AAAR96232 Cancer metastasis
40	3099	99.9	609 17	AAAR94572 Human serum albumi
41	3099	99.9	609 17	AAAR8913 Human serum albumi
42	3099	99.9	609 19	AAAR48095 Human serum albumi
43	3099	99.9	609 20	AAAY06994 Human albumin. Ho
44	3099	99.9	609 22	AAAB04148 Myosin light chain
45	3099	99.9	609 23	ABG32802 Human serum albumi

ALIGNMENTS

RESULT 1

AAAP90388
ID AAP90388 standard; protein: 585 AA.
XX
AC AAP90388;
XX
DT 25-MAR-2003 (updated)
DT 01-NOV-1989 (first entry)
XX
DE Mature human serum albumin polypeptide.
XX
KW Human serum albumin; mature protein; new polypeptides;
KW plasma expanders.
XX
OS Homo sapiens (Human).
XX
PN EP322094-A.
XX
PD 28-JUN-1989.
XX
PF 25-OCT-1988; 88EP-0310000.
XX
PR 30-OCT-1987; 87GB-0025529.
XX
(DELZ) DELTA BIOTECHNOLOGY LTD.
XX
PI Ballance DJ, Hinchliffe E, Geisow MJ, Senior PJ;
XX
DR WPI, 1989-186464/26.
XX
DR N-PSDB; AAAN90128.
XX
PT New N-terminal fragments of human serum albumin


```

XX
DT 25-MAR-2003 (updated)
DT 16-APR-1991 (first entry)
XX
DE Human serum albumin.
XX
KW HSA; folding; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 123..303
FT /label= A
FT Region 1..303
FT /label= B
FT Region 123..585
FT /label= C
XX
PN JP02227079-A.
XX
PD 10-SEP-1990.
XX
PF 25-AUG-1989; 89JP-0217540.
XX
PR 25-AUG-1989; 89JP-0217540.
XX
PA (TOFU ) TONEN CORP.
XX
WPI; 1990-317325/42.
DR N-PSDB; AAQ06099.
XX
PT New human serum albumin fragments - used to bond medicines and for
PT stable folding of protein(s).
XX
PS Claim 1; Fig 8; 24pp; Japanese.
XX
CC Fragments A-C of HSA are expressed as fusion proteins with the
CC signal peptide of E. coli alkaline phosphatase. The fragments are
CC selected for their specific properties. The C-terminal truncated
CC fragment, B, does not bind long-chain fatty acids but does bind to
CC various medicines at the central region. The N-terminal truncated
CC fragment, C, has good stability in protein folding. The central
CC segment, A, has characteristics of both B and C.
CC See also AAQ06099-C06098.
CC (Updated on 25-MAR-2003 to correct PD field.)
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFKDGLGENFKALVLIAPFAQYLQCCPPEDHVKLVNVEFTAKTVADESAE 60
DB 1 DAHSEVAHFKDGLGENFKALVLIAPFAQYLQCCPPEDHVKLVNVEFTAKTVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLYETARRHPYFAPPELLFFAKRYKAATECCQAADKAACLLP 180
DB 121 DVMCTAFHDNEETFLKKYLYETARRHPYFAPPELLFFAKRYKAATECCQAADKAACLLP 180
QY 181 KIDELRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFYSKLVTDLTK 240
DB 181 KIDELRDEGKASSAKORLKCASLOKFGERAFKAWAVARLSORFPKAEFYSKLVTDLTK 240
QY 241 VHTCECHGDLLEACDADRADIKYICENODSISSSKKECCCKPILLEKSHCIAEVENDMPA 300
DB 241 VHTCECHGDLLEACDADRADIKYICENODSISSSKKECCCKPILLEKSHCIAEVENDMPA 300

```

```

QY 301 DLPSLAADFVESKDVCKKNYAERAKDVFGLMFLEYEARHPDYSVLLLRKAKTYETLEKC 360
DB 301 DLPSLAADFVESKDVCKKNYAERAKDVFGLMFLEYEARHPDYSVLLLRKAKTYETLEKC 360
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCFLPQLGEYKFNALLVRYTKVPQVST 420
DB 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCFLPQLGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKCKPEAKRMPCAEADYLSVNLQCVLHKEKTPVSDRYTKCCTES 480
DB 421 PTLVEVSRNLGKVGSKCKCKPEAKRMPCAEADYLSVNLQCVLHKEKTPVSDRYTKCCTES 480
QY 481 LVNRRPCFSALEVDYETVYKPEFNAETFFHADICTLSEKERQIKQALVELVHKPKAT 540
DB 481 LVNRRPCFSALEVDYETVYKPEFNAETFFHADICTLSEKERQIKQALVELVHKPKAT 540
QY 541 KEOLKAVMDDEFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585
DB 541 KEOLKAVMDDEFAAFVEKCKCKADDDKTCFAEEGKKLVAASQAALGL 585

```

RESULT 4

AAR80301
ID AAR80301 standard; Protein; 585 AA.

XX AC AAR80301;

XX XX

DT 25-MAR-2003 (updated)

DT 17-JAN-1996 (first entry)

XX XX

DE Human serum albumin.

XX XX

KW Serum albumin; HSA; aspartyl protease-3; Yap3p;

KW Saccharomyces cerevisiae.

XX OS Homo sapiens.

XX PN W09523957-A1.

XX PD 08-SEP-1995.

XX PF 01-MAR-1995; 95WO-GB00434.

XX PR 05-MAR-1994; 94GB-0004270.

XX XX

PA (DELZ) DELTA BIOTECHNOLOGY LTD.

XX PI KerryWilliams SM, Gilbert SC;

XX DR WPI; 1995-320572/41.

DR N-PSDB; AAQ98695.

PT Yeast with reduced levels of aspartyl protease 3 proteolytic

PT activity - used to secrete human albumin without prodn. of the 45

PT KD fragment

XX Example 1; Page 26-28; 50pp; English.

XX CC The cDNA given in AAQ98695, which encodes HSA (AAR80301), was subjected
CC to site-directed mutagenesis to investigate the role of
CC endoproteases in the generation of a 45 kDa albumin fragment obtd.
CC when the cDNA is expressed in S. cerevisiae. Mutations were: R410A;
CC L407A, L408V, V409A; and R410A, K413Q, K414Q. The latter set of
CC mutations, especially, improved stability of HSA to yeast Yap3p
CC proteolytic cleavage, allowing increased prodn. of recombinant HSA.
CC (Updated on 25-MAR-2003 to correct PI field.)

XX SQ Sequence 585 AA;

Query Match

Best Local Similarity 100.0%; Score 3103; DB 16; Length 585;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHREKDLGEENFKALVLIAPAOYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 60
 Db 1 DAHSEVAHREKDLGEENFKALVLIAPAOYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 60
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREPV 120
 Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREPV 120
 QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFTAKRYKAAFTCCQAAADKAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFTAKRYKAAFTCCQAAADKAACLLP 180
 QY 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLCCKECPLEKSHCIAEVENDEMPA 300
 Db 241 VITECHGDLLECADRADLAKYICENQDSISSKLCCKECPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVSKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360
 Db 301 DLPSLAADFVSKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360
 QY 361 CAADPHCEYAKVDFEKPVLVEPQNLKONCELFQGLGYKFKONALLVRYTKKPVQYST 420
 Db 361 CAADPHCEYAKVDFEKPVLVEPQNLKONCELFQGLGYKFKONALLVRYTKKPVQYST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKEROIKKOTALVELVKKPKPAT 540
 Db 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKEROIKKOTALVELVKKPKPAT 540
 QY 541 KBQLKAVMDDFAAAFVEKCKADDKETCFPAEKGKLIYAASQAALGL 585
 Db 541 KBQLKAVMDDFAAAFVEKCKADDKETCFPAEKGKLIYAASQAALGL 585

RESULT 5

AAO20111
 ID AAO20111 standard; Protein; 585 AA.
 AC AAO20111;
 XX
 DT 06-AUG-2002 (first entry)
 XX
 DE HSA protein sequence related to the growth hormone protein.
 XX
 KW Serum albumin-growth hormone fusion protein; growth hormone;
 XX Down's syndrome.
 OS Unidentified.
 XX
 FN KR99076789-A.
 XX
 PD 15-OCT-1999.
 XX
 PF 25-JUN-1998; 98KR-0704914.
 XX
 PR 30-DEC-1995; 95GB-0026733.
 PR 19-DEC-1996; 96WO-GB03164.
 XX
 PA (DELTA) DELTA BIOTECHNOLOGY LTD.
 XX
 DR WPI; 1997-363680/55.
 DR N-PSDB; AAK99568.
 XX
 PT Serum albumin-growth hormone fusion protein - useful to treat growth
 PT hormone related diseases, e.g. Down's syndrome

XX Disclosure; Fig 6; 2lpp; Korean.
 XX
 CC The invention relates to a serum albumin-growth hormone fusion protein -
 CC useful to treat growth hormone related diseases such as Down's syndrome.
 CC This sequence represents a HSA protein related to the serum albumin-
 CC growth hormone protein of the invention.
 XX
 SQ Sequence 585 AA;
 Query Match 100.0%; Score 3103; DB 18; Length 585;
 Best Local Similarity 100.0%; Pred. No. 9.5e-255;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHREKDLGEENFKALVLIAPAOYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 60
 Db 1 DAHSEVAHREKDLGEENFKALVLIAPAOYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 60
 QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREPV 120
 Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPPLRLVREPV 120
 QY 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFTAKRYKAAFTCCQAAADKAACLLP 180
 Db 121 DVNCTAFHDNEETFLKKYLVEIARRHPYFYAPPELLFTAKRYKAAFTCCQAAADKAACLLP 180
 QY 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240
 Db 181 KLDELDRSGKASSAKQRLKASLOKFGERAFKANAVARLSORFPKAEFAEVSKLVTDLTK 240
 QY 241 VITECHGDLLECADRADLAKYICENQDSISSKLCCKECPLEKSHCIAEVENDEMPA 300
 Db 241 VITECHGDLLECADRADLAKYICENQDSISSKLCCKECPLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVSKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360
 Db 301 DLPSLAADFVSKDVCKNYAEAKDVFGLMFLYEYARRHPDYSVVLRLRLAKTYETLEKC 360
 QY 361 CAADPHCEYAKVDFEKPVLVEPQNLKONCELFQGLGYKFKONALLVRYTKKPVQYST 420
 Db 361 CAADPHCEYAKVDFEKPVLVEPQNLKONCELFQGLGYKFKONALLVRYTKKPVQYST 420
 QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 Db 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 QY 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKEROIKKOTALVELVKKPKPAT 540
 Db 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKEROIKKOTALVELVKKPKPAT 540
 QY 541 KBQLKAVMDDFAAAFVEKCKADDKETCFPAEKGKLIYAASQAALGL 585
 Db 541 KBQLKAVMDDFAAAFVEKCKADDKETCFPAEKGKLIYAASQAALGL 585

RESULT 6

AAO20111
 ID AAO20111 standard; protein; 585 AA.
 AC AAO20111;
 XX
 DT 08-AUG-2000 (first entry)
 XX
 DE Amino acid sequence of a human albumin protein.
 XX
 KW Human; albumin; ischemic state; serum protein; metal ion salt;
 KW perioperative ischemia; ischemia; myocardial infarction;
 KW progressive coronary artery disease.
 XX
 OS Homo sapiens.
 XX
 FN Key Location/Qualifiers
 FT Modified-site 1

PT /note= "optionally acetylated, and claimed under
FT claim 56"
PN WO200020840-A1.
XX 13-APR-2000.
PD 01-OCT-1999; 99WO-US22905.
XX 02-OCT-1998; 98US-0102738.
XX 02-OCT-1998; 98US-0165581.
XX 02-OCT-1998; 98US-0185926.
XX 11-JAN-1999; 99US-0115392.
XX (ISCH-) ISCHEMIA TECHNOLOGIES INC.
XX Bar-Or D, Lau E, Winkler JV;
XX WPI; 2000-303843/26.
XX New method for the continuous detection of ischemic states comprises
PT detecting and quantifying the existence of an alteration of the serum
PT protein albumin -
XX Disclosure; Page 97-100; 105pp; English.
XX The present sequence represents human albumin protein. The specification
CC describes a method for the continuous detection of ischemic states. The
CC method comprises detecting and quantifying the existence of an alteration
CC of the serum protein albumin. The method comprises contacting a
CC biological sample containing albumin from the patient with an excess
CC quantity of a metal ion salt, where the metal ion binds to the N-terminus
CC of naturally occurring human albumin, to form a mixture containing bound
CC metal ions and unbound metal ions, and then determining the amount of
CC metal ions bound to the albumin N-terminus. The amount of bound metal
CC ions is correlated to a known value to determine the occurrence or
CC non-occurrence of an ischemic event. The methods are useful for detection
CC of ischemic states. The methods are also useful for distinguishing
CC pericardial ischemia from ischemia caused by , amongst other things,
CC myocardial infarctions and progressive coronary artery disease.
XX
XX Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255; Indels 0; Gaps 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQPERNECFLOHDDNPNLPRLVREPV 120
DB 61 NCDKSLHTLFGDKLCTVATRETYGEMADCCAKQPERNECFLOHDDNPNLPRLVREPV 120
QY 121 DVNCTAFHDNETHFKLYLIEIARRHPYFAPPELLFFAKRYKAAFTCCQAAKACLLP 180
DB 121 DVNCTAFHDNETHFKLYLIEIARRHPYFAPPELLFFAKRYKAAFTCCQAAKACLLP 180
QY 181 KIDELDEKASSAKQRLKASLOKXGERAFKAWAVARLSORFPKAEFAVSKLYTDTLK 240
DB 181 KIDELDEKASSAKQRLKASLOKXGERAFKAWAVARLSORFPKAEFAVSKLYTDTLK 240
QY 241 VHTECCHGDLLECDRADLAKYICENQDISISKLKECEKPLEKSHCIAEVENDEMPA 300
DB 241 VHTECCHGDLLECDRADLAKYICENQDISISKLKECEKPLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADPVESKVCNKAFAKDFLGMFLYEAARRHPDYVSVLLLRILAKYVETLEKC 360
DB 301 DLPSLAADPVESKVCNKAFAKDFLGMFLYEAARRHPDYVSVLLLRILAKYVETLEKC 360
QY 361 CAADPHECVAKVDFEFPQNLKONCELFEQGEYKFNALVRYTKVPQVST 420

DB 361 CAADPHECVAKVDFEFPQNLKONCELFEQGEYKFNALVRYTKVPQVST 420
QY 421 PTLVEVSRNLGVGSKCKHPEAKRMPCABDYLSVNLQVLHETKTPVSDRVTKCTES 480
DB 421 PTLVEVSRNLGVGSKCKHPEAKRMPCABDYLSVNLQVLHETKTPVSDRVTKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKBFNAETFTFHADICTLSEKQIKKQATLVELVKKHPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKBFNAETFTFHADICTLSEKQIKKQATLVELVKKHPKAT 540
QY 541 KEQLKAVMDDFRAAFVEKCKCKADDETCFAEEGKKLVASQAALGL 585
DB 541 KEQLKAVMDDFRAAFVEKCKCKADDETCFAEEGKKLVASQAALGL 585
RESULT 7
AAV83946
ID AAY83946 standard; Protein; 585 AA.
XX AAY83946;
AC AAY83946;
XX 28-JUL-2000 (first entry)
DT 28-JUL-2000 (first entry)
XX Yeast codon-biased recombinant human serum albumin protein.
DE Recombinant; human serum albumin; HSA; yeast codon bias; host cell;
XX overlapping oligonucleotide; expression vector.
KW Homo sapiens.
OS Synthetic.
XX CNL239103-A.
XX 22-DEC-1999.
XX 17-JUN-1998; 98CN-0102506.
XX 17-JUN-1998; 98CN-0102506.
XX (HALJ-) HALJI BIOENGINEERING CO LTD.
XX Li S, Lu D;
XX WPI; 2000-351198/31.
XX N-PSDB; AAAL0091.
XX Process for preparing recombinant human serum albumin - which comprises
PT yeast biased sex codons
XX Disclosure; Fig 1; 44pp; Chinese.
XX The method relates to a method of recombinantly producing human serum
CC albumin (HSA) in yeast by altering the coding sequence of HSA to
CC comprise a yeast codon bias. The complete HSA gene (AAAL0091) was
CC generated as three synthetic fragments (AAAL0092-A10094) joined by
CC recombinant DNA technology. Each HSA fragment was synthesised from
CC overlapping oligonucleotide fragments that were extended. This sequence
CC represents the complete sequence of the HSA encoded by the human gene
CC with a yeast codon bias. The invention also covers a recombinant
CC expression vector, yeast host cells carrying the recombinant expression
CC vector and the process for producing human serum albumin in the yeast
CC host cell, especially in secretory mode.
XX Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 21; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255; Indels 0; Gaps 0;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHFKDLGEENFKALVLIAPAYLQCCPFEDHVKLVNEVTEFAKTCVADESAAE 60

```

QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPRVPEV 120
Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPRVPEV 120
QY 121 DVMCTAFHDNEETFLKKLYLIEARHPYFFAPPELLFAKRYKAATECCQADKAACILP 180
Db 121 DVMCTAFHDNEETFLKKLYLIEARHPYFFAPPELLFAKRYKAATECCQADKAACILP 180
QY 181 KLDELREGKASSAKORLKCSAQKFGERAFKAWAVARLSORFPKAEFAEYSKLVTDLT 240
Db 181 KLDELREGKASSAKORLKCSAQKFGERAFKAWAVARLSORFPKAEFAEYSKLVTDLT 240
QY 241 VHTPECCGDLLECADRDADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTPECCGDLLECADRDADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNFAEAKDVFGLGMFLYFARHPDYISVLLLLRLAKTYETTLKC 360
Db 301 DLPSLAADFVESKDVCKNFAEAKDVFGLGMFLYFARHPDYISVLLLLRLAKTYETTLKC 360
QY 361 CAADAPHECYAKVDFEFLVEEPONLIKQNCLEFQGLGEYKFQNALVRYTKKVPQVST 420
Db 361 CAADAPHECYAKVDFEFLVEEPONLIKQNCLEFQGLGEYKFQNALVRYTKKVPQVST 420
QY 421 PTLVEVSRLNGKVGSKCKHPFAKMPCAEDYLSVLNOLCVLHEKTPVSDRVTKKCTES 480
Db 421 PTLVEVSRLNGKVGSKCKHPFAKMPCAEDYLSVLNOLCVLHEKTPVSDRVTKKCTES 480
QY 481 LVNRRPCFSALEVDIYVYKFNATFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDIYVYKFNATFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDFAAFVEKCKADKDKTCFAEKGKLVAAQAALGL 585
Db 541 KEQLKAVMDFAAFVEKCKADKDKTCFAEKGKLVAAQAALGL 585

RESULT 8
ABBY79006
ID ID ABBY79006 standard; Protein; 585 AA.
XX AC ABBY79006;
XX DT 01-AUG-2002 (first entry)
XX DE Human mature albumin protein SEQ ID NO:18.
XX KW Human; growth hormone; hGH; albumin; human serum albumin; HSA;
KW albumin fusion protein; cytostatic; anorectic; immunosuppressive;
KW antidiabetic; antirheumatic; antiarthritic; psoriatic; cancer;
KW non-Hodgkin's lymphoma; obesity; type I diabetes mellitus; rheumatoid arthritis.
XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT 1..194
XX FT /label= 1
XX FT Domain
XX FT 1..105
XX FT /label= subdomain
XX FT Disulfide-bond 53..62
XX FT Disulfide-bond 75..91
XX FT Disulfide-bond 90..101
XX FT Region 106..119
XX FT /note= "flexible inter-subdomain linker region"
XX FT Domain 120..194
XX FT /label= subdomain
XX FT Disulfide-bond 124..169
XX FT Disulfide-bond 168..177
XX FT Domain 195..387
XX FT /label= 2
XX FT Domain 195..291
XX FT /label= subdomain

```

```

FT Disulfide-bond 245..253
FT Disulfide-bond 265..279
FT Disulfide-bond 278..289
FT Region
FT /note= "flexible inter-subdomain linker region"
FT Domain 316..387
FT /label= subdomain
FT Disulfide-bond 316..361
FT Disulfide-bond 360..369
FT Domain 388..585
FT /label= 3
FT Domain 388..491
FT /label= subdomain
FT Disulfide-bond 392..438
FT Disulfide-bond 437..448
FT Disulfide-bond 461..477
FT Disulfide-bond 476..487
FT Region 492..511
FT /note= "flexible inter-subdomain linker region"
FT Domain 512..585
FT /label= subdomain
FT Disulfide-bond 514..559
FT Disulfide-bond 558..567
XX WO200179442-A2.
XX PD 25-OCT-2001.
XX PF 12-APR-2001; 2001WO-US11850.
XX PR 12-APR-2000; 2000US-229358P.
XX PR 25-APR-2000; 2000US-199384P.
XX PR 21-DEC-2000; 2000US-256931P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX FI Rosen CA, Haseltine WA;
XX DR WPI; 2001-611723/70.
XX DR N-PSDB; ABBY7288.
XX FT New albumin fusion proteins, useful for treating diseases and disorders
XX FT such as cancer, comprise therapeutic protein fused to albumin -
XX PS Claim 1; Fig 11; 413pp; English.
XX CC The present invention describes an albumin fusion protein (I) comprising
XX CC a therapeutic protein: X and (a fragment or variant of) albumin
XX CC comprising a the fully defined sequence in ABBY79006 of 585 amino acids,
XX CC (where the fragment or variant has albumin or therapeutic protein: X
XX CC activity). (I) can have cytostatic, anorectic, immunosuppressive,
XX CC antidiabetic, antirheumatic, antiarthritic and psoriatic activities.
XX CC Albumin fusion proteins are stabilised therapeutic proteins e.g.
XX CC antibodies to CS, C242 and C890 useful for treating various diseases
XX CC and disorders such as non-Hodgkin's lymphoma, cancer, obesity,
XX CC transplant rejection, type I diabetes mellitus, rheumatoid arthritis
XX CC and psoriasis. Fusing albumin to therapeutic proteins stabilises the
XX CC therapeutic protein, extends the shelf life and retains the in vitro or
XX CC in vivo biological activity. It also reduces the need to formulate
XX CC protein solutions with large excesses of carrier proteins to prevent
XX CC loss of therapeutic proteins due to factors such as binding to the
XX CC container. The fusion proteins are easily dispensed with a simple
XX CC formulation requiring minimal post storage manipulation. The fusion of
XX CC therapeutic proteins to albumin confers stability in aqueous or other
XX CC solution. The present sequence represents the mature human albumin (HA)
XX CC protein which is used in the exemplification of the present invention.
XX SQ Sequence 585 AA;
XX Query Match 100.0%; Score 3103; DB 22; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDEHVKLVNVEVTEFAKTCVADESSE 60
 Db 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDEHVKLVNVEVTEFAKTCVADESSE 60
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLYRPV 120
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLYRPV 120
 QY 121 DVMCTAFHDNEETFLKKYLFIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAACLLP 180
 Db 121 DVMCTAFHDNEETFLKKYLFIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAACLLP 180
 QY 181 KLDELDEGKASAKORLKCSLOKFGERAFKAWARLSQRFKAEFAVSKLYVDLTK 240
 Db 181 KLDELDEGKASAKORLKCSLOKFGERAFKAWARLSQRFKAEFAVSKLYVDLTK 240
 QY 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKCECCERPLEKSHCHIAEYENDEMPA 300
 Db 241 VHTCCGDLLECCADRADLAKYICENQDSISSKLKCECCERPLEKSHCHIAEYENDEMPA 300
 QY 301 DLPISLAADFVESKDVCKNYAKADVFTLGMFLYIYARRHPDYSVLLRLAKTETYLEKC 360
 Db 301 DLPISLAADFVESKDVCKNYAKADVFTLGMFLYIYARRHPDYSVLLRLAKTETYLEKC 360
 QY 361 CAADAPHECYAKVDFEPLVEEPQNLKONCELPOLGEYKFNQALIVRYTKKVPQVST 420
 Db 361 CAADAPHECYAKVDFEPLVEEPQNLKONCELPOLGEYKFNQALIVRYTKKVPQVST 420
 QY 421 PTLVEVSRNLGVSKCKHPKAKMPCAEEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
 Db 421 PTLVEVSRNLGVSKCKHPKAKMPCAEEDYLSVLNQLCVLHEKTPVSDRVTKCTES 480
 QY 481 LVNRPFCFSALEVDIYVPKEFNAETFTFHADICTLSEKEROIKQATLVELVKHFKPAT 540
 Db 481 LVNRPFCFSALEVDIYVPKEFNAETFTFHADICTLSEKEROIKQATLVELVKHFKPAT 540
 QY 541 KEQLKAMDDFAAFVEKCKADKDETCTFAEKGKLVAAASQAALGL 585
 Db 541 KEQLKAMDDFAAFVEKCKADKDETCTFAEKGKLVAAASQAALGL 585

RESULT 9

AAE13399

ID AAE13399 standard; Protein; 585 AA.

AC AAE13399;

XX 12-FEB-2002 (first entry)

DE Human albumin (HA) protein.

XX Human; albumin; HA; fusion protein; immune system disorder; syphilis;
 KW transplant rejection; blood related disorder; myocardial infarction;
 KW hyperproliferative disorder; acute myeloid leukaemia; renal disorder;
 KW glomerulonephritis; cardiovascular disorder; arrhythmia; rhinitis;
 KW respiratory disorder; neurological disease; Alzheimer's disease;
 KW endocrine disorder; pheochromocytoma; reproductive system disorder;
 KW measles; gastrointestinal disorder; irritable bowel syndrome; HIV;
 KW human immunodeficiency virus; wound healing; renal cell carcinoma;
 KW melanoma; gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FH Domain 54..61

FT /label= Loop_I

FT 76..89

FT /label= Loop_II

FT 92..100

FT /label= Loop_III

FT 170..176

FT /label= Loop_IV

FT 247..252

FT Domain /label= Loop_V
 266..277
 FT /label= Loop_VI
 280..288
 FT Domain /label= Loop_VII
 362..368
 FT Domain /label= Loop_VIII
 439..447
 FT Domain /label= Loop_IX
 461..475
 FT Domain /label= Loop_X
 478..486
 FT Domain /label= Loop_XI
 560..566
 FT /label= Loop_XII

XX WO200179258-A1.

PN 25-OCT-2001.

XX 12-APR-2001; 2001WO-US12008.

XX 12-APR-2000; 2000US-229358P.

PR 25-APR-2000; 2000US-199384P.

PR 21-DEC-2000; 2000US-256931P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (PRIN-) PRINCIPIA PHARM CORP.

XX Rosen CA, Sadeghi H, Prior CP, Turner AJ;

XX WPI; 2001-602931/68.

DR N-PSDB; RAD22287.

XX Albumin fusion proteins comprising a therapeutic protein and albumin,

PT useful in the treating metastatic renal cell carcinoma, metastatic

PT melanoma, malignant melanoma, renal cell carcinoma, HIV (human

PT immunodeficiency virus) or infection -

XX Claim 1; Fig 9; 325pp; English.

XX The invention relates to albumin fusion proteins comprising therapeutic

CC protein and human albumin (HA). The albumin fusion proteins are useful

CC in the treatment, prevention, diagnosis, and/or detection of diseases,

CC disorders such as immune system disorders (transplant rejection); blood

CC related disorders (myocardial infarction); hyperproliferative disorders

CC (childhood acute myeloid leukaemia); renal disorder (glomerulonephritis);

CC cardiovascular disorders (arrhythmias); neurological diseases (Alzheimer's disease);

CC endocrine disorders (pheochromocytoma); reproductive system disorders

CC (syphilis); infectious diseases (measles); gastrointestinal disorders

CC (irritable bowel syndrome) and wound healing. The albumin fusion

CC proteins are also used in the treatment of metastatic renal cell

CC carcinoma, metastatic melanoma, malignant melanoma and HIV (human

CC immunodeficiency virus) infection. Nucleic acid encoding albumin fusion

CC protein is useful in gene therapy. The present sequence is human

CC albumin (HA) protein.

XX Sequence 585 AA;

SQ

Query Match 100.0%; Score 3103; DB 22; Length 585;

Best Local Similarity 100.0%; Pred. No. 9.5e-255;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDEHVKLVNVEVTEFAKTCVADESSE 60

Db 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDEHVKLVNVEVTEFAKTCVADESSE 60

QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLYRPV 120

Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNDNPLRLYRPV 120

QY 121 DVMCTAFHDNEETFLKKYLFIARRHPYFYAPPELLFAKRYKAAFTCCQAADKAACLLP 180

```

Db 121 DVMCTAFDNEETFLKKYLYEIARRHPYFAPELLFFAKRYAAATECCQADRAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEYSKLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEYSKLVTDLT 240
QY 241 VHTCCGDLLECADRADLAKYICENODSISSSKKECCCKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECADRADLAKYICENODSISSSKKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNAAEKDVLGMFLYIYARRHPDYSVLLLRKAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNAAEKDVLGMFLYIYARRHPDYSVLLLRKAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585

RESULT 10
AA052567
ID AA052567 standard; Protein; 585 AA.
XX
AC AA052567;
XX
DT 05-FEB-2002 (first entry)
XX
DE Mature human serum albumin.
XX
KW Human; serum albumin; HA; antiinflammatory; immunosuppressive; cardiant;
KW nootropic; neuroprotective; gene therapy; immune disorder; wound healing;
KW hyperproliferative disorder; renal disorder; cardiovascular disorder;
KW respiratory disorder; neurological disease; endocrine disorder;
KW reproductive system disorder; infectious disease;
KW gastrointestinal disorder.
XX
OS Homo sapiens.
XX
PN WO200179444-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-US12013.
XX
PR 12-APR-2000; 2000US-229358P.
PR 25-APR-2000; 2000US-199384P.
PR 21-DEC-2000; 2000US-256931P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI: 2001-616755/71.
XX
XX N-PSDB; ABA03057.
XX
XX Albumin fusion proteins comprising a therapeutic protein and albumin,
XX useful in the treating immune system disorders (e.g. transplant
XX rejection), blood related disorders (e.g. myocardial infarction) and
XX hyperproliferative disorders
XX

```

```

PS Claim 1: Fig 15; 606pp; English.
XX
CC The present invention relates to albumin fusion proteins, which comprise
CC a therapeutic protein and albumin. The present sequence is the protein
CC sequence for mature human serum albumin (HA), which was used to generate
CC the fusion proteins of the present invention. The albumin fusion proteins
CC are useful in the treatment, prevention, diagnosis, and/or detection of
CC diseases/disorders such as immune system disorders (e.g. myocardial
CC rejection), blood related disorders (e.g. childhood acute myeloid leukemia),
CC hyperproliferative disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
CC renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g.
CC arrhythmias), respiratory disorders (e.g. non-allergic rhinitis),
CC neurological diseases (e.g. Alzheimer's disease), endocrine disorders
CC (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis),
CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.
CC irritable bowel syndrome) and wound healing.
XX
SQ Sequence 585 AA;
Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHREFKOLGGEENFKALVIAFAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESAE 60
Db 1 DAHSEVAHREFKOLGGEENFKALVIAFAQYLOQCPFFEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPPLRVLPREV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPPLRVLPREV 120
QY 121 DVMCTAFDNEETFLKKYLYEIARRHPYFAPELLFFAKRYAAATECCQADRAACLLP 180
Db 121 DVMCTAFDNEETFLKKYLYEIARRHPYFAPELLFFAKRYAAATECCQADRAACLLP 180
QY 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEYSKLVTDLT 240
Db 181 KLDELDEGKASSAKQRLKASLOKFGGERAFKAWAVARLSQRPFAEFAEYSKLVTDLT 240
QY 241 VHTCCGDLLECADRADLAKYICENODSISSSKKECCCKPLLEKSHCIAEVENDEMPA 300
Db 241 VHTCCGDLLECADRADLAKYICENODSISSSKKECCCKPLLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNAAEKDVLGMFLYIYARRHPDYSVLLLRKAKTYETTLK 360
Db 301 DLPSLAADFVESKDVCKNAAEKDVLGMFLYIYARRHPDYSVLLLRKAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
Db 361 CAADPHECYAKVDFEFKPLVEEPONLIKQNCLEFQELGEYKFNALLVRYTKVPQVST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKRMPCAEDYLSVLNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDYTVPKFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT 540
QY 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFAVEKCKCKADDDKTCFAEFGKKLVAASQAALGL 585

RESULT 11
AA013129
ID AA013129 standard; Protein; 585 AA.
XX
AC AA013129;
XX
DT 28-JAN-2002 (first entry)
XX
DE Human albumin (HA).

```


xx Human; albumin; HA; fusion protein; therapeutic protein; vulnery;

xx immune system disorder; transplant rejection; blood related disorder;

xx myocardial infarction; hyperproliferative disorder; glomerulonephritis;

xx childhood acute myeloid leukemia; cardiovascular disorder; arrhythmia;

xx respiratory disorder; gene therapy; non-allergic rhinitis; nontropic;

xx neurological disease; Alzheimer's disease; reproductive system disorder;

xx endocrine disorder; pheochromocytoma; infectious disease; antiarthritic;

xx measles; gastrointestinal disorder; irritable bowel syndrome; syphilis;

xx wound healing; antiinflammatory; immunosuppressive; neuroprotective;

xx cardiant; cytostatic; anti-leukemic; antirheumatic; antimicrobial;

xx renal disorder.

xx Homo sapiens.

xx Key Location/Qualifiers

xx Domain 54..61 /label= Loop_I

xx Domain 76..89 /label= Loop_II

xx Domain 92..100 /label= Loop_III

xx Domain 170..176 /label= Loop_IV

xx Domain 247..252 /label= Loop_V

xx Domain 266..277 /label= Loop_VI

xx Domain 280..288 /label= Loop_VII

xx Domain 362..368 /label= Loop_VIII

xx Domain 439..447 /label= Loop_IX

xx Domain 461..475 /label= Loop_X

xx Domain 478..486 /label= Loop_XI

xx Domain 560..566 /label= Loop_XII

xx WO200179443-A2.

xx 25-OCT-2001.

xx 12-APR-2001; 2001WO-US11924.

xx 12-APR-2000; 2000US-229358P.

xx 25-APR-2000; 2000US-199384P.

xx 21-DEC-2000; 2000US-256931P.

xx (HUMA-) HUMAN GENOME SCI INC.

xx Rosen CA, Haseltine WA;

xx WPI; 2001-616754/71.

xx N-PSDB; AAD21638.

xx Albumin fusion proteins comprising a therapeutic protein and albumin,

xx useful in the treating immune system disorders (e.g. transplant

xx rejection), blood related disorders (e.g. myocardial infarction) and

xx hyperproliferative disorders -

xx Claim 1; Fig 9; 38Opp; English.

xx The invention relates to albumin fusion proteins comprising therapeutic

xx protein and human albumin (HA). Therapeutic protein fused to albumin

xx have an extended shelf-life. The albumin fusion proteins are useful in

xx the treatment, prevention, diagnosis and/or detection of diseases,

xx disorders such as immune system disorders (e.g. transplant rejection),

xx blood related disorders (e.g. myocardial infarction), hyperproliferative

xx disorders (e.g. childhood acute myeloid leukaemia), renal disorders

xx (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias),

CC respiratory disorders (e.g. non-allergic rhinitis), neurological

CC diseases (e.g. Alzheimer's disease), endocrine disorders (e.g.

CC pheochromocytoma), reproductive system disorders (e.g. syphilis),

CC infectious diseases (e.g. measles), gastrointestinal disorders (e.g.

CC irritable bowel syndrome) and wound healing. Nucleic acids encoding

CC albumin fusion protein is used in gene therapy. The present sequence

CC is human albumin (HA) protein.

xx Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;

Best Local Similarity 100.0%; Pred. No. 9,5e-255;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFRALVLIAPAYLOOCFFEDHVKLVNEVTEFAKTCVADESAAE 60

DB 1 DAHSEVAHFRKDLGEENFRALVLIAPAYLOOCFFEDHVKLVNEVTEFAKTCVADESAAE 60

QY 61 NCDKSLHTLFGDKICTVATLRETYGEMADCCAKQPERNECFLOHDDNPNLPRVREPV 120

DB 61 NCDKSLHTLFGDKICTVATLRETYGEMADCCAKQPERNECFLOHDDNPNLPRVREPV 120

QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180

DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFAPPELLFFAKRYKAAFTCCQADKAAACLLP 180

QY 181 KLDELNDGKASSAKORLKCASQKQGERAFRAKAVARIQSFFPKAFPAVSKLVTDLTG 240

DB 181 KLDELNDGKASSAKORLKCASQKQGERAFRAKAVARIQSFFPKAFPAVSKLVTDLTG 240

QY 241 VHTCCGDLLECCADRADLAKYICENQDSISKLECEKPLEKSHCHIAEYENDEMPA 300

DB 241 VHTCCGDLLECCADRADLAKYICENQDSISKLECEKPLEKSHCHIAEYENDEMPA 300

QY 301 DLPFLAADFVESKQCKNYAEAKDVFGLMFLYFARRHPDYSVVLRLAKTYETTTLEKC 360

DB 301 DLPFLAADFVESKQCKNYAEAKDVFGLMFLYFARRHPDYSVVLRLAKTYETTTLEKC 360

QY 361 CAADPHECYAKVDFEFPNLIKONCELFQGEYKFNALLVYTKVPQVST 420

DB 361 CAADPHECYAKVDFEFPNLIKONCELFQGEYKFNALLVYTKVPQVST 420

QY 421 PTLVEYSRLNGVSKCKHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480

DB 421 PTLVEYSRLNGVSKCKHPEAKRMPCAEYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480

QY 481 LVNRRPCFSALEVDYETYPKEFNAETFFHADICTLSEKEQIKKQALVELVKHKKPAT 540

DB 481 LVNRRPCFSALEVDYETYPKEFNAETFFHADICTLSEKEQIKKQALVELVKHKKPAT 540

QY 541 KEQLKAVMDFAAFVEKCKKADKTCFAEKGKLVAAASQAALGL 585

DB 541 KEQLKAVMDFAAFVEKCKKADKTCFAEKGKLVAAASQAALGL 585

RESULT 12

AAE12403

ID AAE12403 standard; Protein; 585 AA.

XX AAE12403;

AC AAE12403;

XX 18-DEC-2001 (first entry)

DT 18-DEC-2001 (first entry)

XX Human albumin (HA).

DE Human; albumin; HA; immune system disorder; transplant rejection;

xx blood related disorder; myocardial infarction; glomerulonephritis;

xx hyperproliferative disorder; childhood acute myeloid leukaemia;

xx renal cell carcinoma; cardiovascular disorder; vulnery; melanoma;

xx arrhythmia; respiratory disorder; non-allergic rhinitis; anti-leukemic;

xx neurological disease; Alzheimer's disease; endocrine disorder; measles;

xx pheochromocytoma; reproductive system disorder; neuroprotective; syphilis;

xx infectious disease; gastrointestinal disorder; wound healing; nontropic;

irritable bowel syndrome; HIV; human immunodeficiency virus infection; cytostatic; antiinflammatory; gene therapy; immunosuppressive; cardiac; antiarthritic; antirheumatic; renal disorder; antimicrobial.

Homo sapiens.

Key Location/Qualifiers
Domain 54..61
/label= Loop_I
Domain 76..89
/label= Loop_II
Domain 92..100
/label= Loop_III
Domain 170..176
/label= Loop_IV
Domain 247..252
/label= Loop_V
Domain 266..277
/label= Loop_VI
Domain 280..288
/label= Loop_VII
Domain 362..368
/label= Loop_VIII
Domain 439..447
/label= Loop_IX
Domain 461..475
/label= Loop_X
Domain 478..486
/label= Loop_XI
Domain 560..566
/label= Loop_XII

WO2001179480-A1.

25-OCT-2001.

12-APR-2001; 2001WO-US11991.

12-APR-2000; 2000US-229358P.

25-APR-2000; 2000US-199384P.

21-DEC-2000; 2000US-256931P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Haseltine WA;

WPI: 2001-616756/71.

N-PSDB; AAD20005.

Albumin fusion proteins comprising a therapeutic protein and albumin, useful in the treating metastatic renal cell carcinoma, metastatic melanoma, malignant melanoma, renal cell carcinoma, HIV (human immunodeficiency virus) or infection -

Claim 1; Fig 9; 394pp; English.

The invention relates to human albumin (HA) fusion proteins and their corresponding nucleic acid sequences. Therapeutic proteins fused to albumin or its fragments have an extended shelf-life. The albumin fusion proteins are useful in the treatment, prevention, diagnosis, and/or detection of diseases, disorders such as immune system disorders (e.g. transplant rejection), blood related disorders (e.g. myocardial infarction), hyperproliferative disorders (e.g. childhood acute myeloid leukemia, metastatic renal cell carcinoma, metastatic melanoma, malignant melanoma, renal cell carcinoma), renal disorders (e.g. glomerulonephritis), cardiovascular disorders (e.g. arrhythmias), respiratory disorders (e.g. non-allergic rhinitis), neurological diseases (e.g. Alzheimer's disease), endocrine disorders (e.g. pheochromocytoma), reproductive system disorders (e.g. syphilis), infectious diseases (e.g. measles), gastrointestinal disorders (e.g. irritable bowel syndrome), HIV (human immunodeficiency virus) infection and wound healing. Nucleic acids encoding albumin fusion protein is used in gene therapy. The present sequence is human albumin.

XX	Sequence	585 AA;	Query Match	100.0%;	Score 3103;	DB 22;	Length 585;	
SQ	Best Local Similarity	100.0%;	Pred. No. 9.5e-255;	Matches 585;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNVEVTEFAKTCVADESAE	60					
Db	1	DAHKEVAHRFKDLGEENFKALVLIAPAYLQOCPPEDHVKLVNVEVTEFAKTCVADESAE	60					
QY	61	NCKDSLHTLFGDKLCITVATIRETYGEMADCCAKQEPERNECFLOHKDDNPMLPLVRPEV	120					
Db	61	NCKDSLHTLFGDKLCITVATIRETYGEMADCCAKQEPERNECFLOHKDDNPMLPLVRPEV	120					
QY	121	DVMCTAFHDNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAFTCCQADRAACAACLLP	180					
Db	121	DVMCTAFHDNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAFTCCQADRAACAACLLP	180					
QY	181	KLDELREGKASSAKORLKCASIQKFGGERAFKAWARLSORPPKAEFAEYKSLVTDLT	240					
Db	181	KLDELREGKASSAKORLKCASIQKFGGERAFKAWARLSORPPKAEFAEYKSLVTDLT	240					
QY	241	VHTECHGDLLESCADRADLAKYICENQDSISSKLECCCKPLEKSHCIAFVENDEMPA	300					
Db	241	VHTECHGDLLESCADRADLAKYICENQDSISSKLECCCKPLEKSHCIAFVENDEMPA	300					
QY	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETLEKC	360					
Db	301	DLPSLAADFVESKDVCKNYAEAKDVLGMFLYFYARRHPDYSVLLRLAKTYETLEKC	360					
QY	361	CAAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFQNALIVRTTKVPQVST	420					
Db	361	CAAADPHECYAKVDFEFKPLVEEPQNLKQNCSELFQOLGEYKFQNALIVRTTKVPQVST	420					
QY	421	PTLVEVSRNLGKVGSKCKKHPKAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES	480					
Db	421	PTLVEVSRNLGKVGSKCKKHPKAEADYLSVLNQLCVLHEKTPVSDRVTKCCTES	480					
QY	481	LVNRRPCFSALEVDVYVPKEFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT	540					
Db	481	LVNRRPCFSALEVDVYVPKEFNAETFFHADICTLSEKERQIKKOTALVELVHKPKAT	540					
QY	541	KEQLKAVMDFFAAFEVKCKCKADDDKTCFAEEGKLVAAQAALGL	585					
Db	541	KEQLKAVMDFFAAFEVKCKCKADDDKTCFAEEGKLVAAQAALGL	585					

RESULT 13

AAE08578

ID AAE08578 standard; Protein; 585 AA.

AC AAE08578;

XX 19-NOV-2001 (first entry)

DT Human serum albumin (HSA).

DE Human; albumin; cancer; cell proliferation; drug screening; biopsy.

KW Homo sapiens.

XX OS6274305-B1.

XX 14-AUG-2001.

XX 19-DEC-1996; 96US-0769746.

XX 19-DEC-1996; 96US-0769746.

XX (TUFT) UNIV TUFTS.

XX Sonnenschein C, Soto AM;

PI

XX WPI: 2001-540371/60.
DR N-PSDB; AAD11488.
XX
PT Measuring human cell proliferation, useful in drug screening to
PT determine the potential for inhibiting cancer cell proliferation and
PT for evaluating biopsied tumors, comprises employing albumin-derived
PT peptide -
XX
XX Claim 1: Fig 1; 20pp; English.
XX
CC The invention related to a method for testing cancer cells. The method is
CC useful for measuring human cancer cell proliferation, particularly for
CC determining the potential for inhibiting cancer cells proliferation using
CC albumin-derived peptides. The invention is also useful for drug screening
CC assays, as well as for evaluating biopsied tumors. The present sequence
CC is human serum albumin (HSA) related to the invention.
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 22; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAFHFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAFHFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDDNPNLPLVRPEV 120
DB 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMTAFHNEETFLKKLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLIP 180
DB 121 DVMTAFHNEETFLKKLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLIP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQFPKAEFAEVS KLIVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQFPKAEFAEVS KLIVTDLTK 240
QY 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMA 300
DB 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCPELFEQGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCPELFEQGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEYSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480
DB 421 PTLVEYSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480
QY 481 LVNRRPCFSALEVDYETVYVPEKFAETTFPHADICTLSEKERQIKKQATALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDYETVYVPEKFAETTFPHADICTLSEKERQIKKQATALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGSKLVAAASQAALGL 585
DB 541 KEQLKAVMDDFAAFEVKCKCKADDKETCFABEGSKLVAAASQAALGL 585

RESULT 14
ABG71291
ID ABG71291 standard; Protein; 585 AA.

XX AC ABG71291;
XX AC
XX DT 08-JAN-2003 (first entry)
XX

DE Glycosylated protein determination associated protein.
XX
KW Standard substance; accuracy control substance; glycosylated protein;
KW glycosylated albumin; fructosamine; diabetes; antidiabetic.
XX
OS Unidentified.
XX
PN JP2002243731-A.
XX
XX 28-AUG-2002.
XX
XX 21-FEB-2001; 2001JP-0045085.
XX
XX 21-FEB-2001; 2001JP-0045085.
XX
PA (KOKU-) KOKUSAI SHIYAKU KK.
PA (YOSH) YOSHITOMI PHARM IND KK.
XX
XX WPI; 2002-744850/81.
XX
XX A standard substance for determination of glycosylated protein
PT including glycosylated albumin and fructosamine, used in diagnosis of
PT diabetes -
XX
XX Disclosure; Page 4; 6pp; Japanese.
XX
CC The present invention relates to a new standard and accuracy control
CC substance for determination of glycosylated protein. The invention is
CC useful for determination of glycosylated protein in the diagnosis of
CC diabetes. Glycosylated albumin and fructosamine provide favourable
CC dilution linearity. The present amino acid sequence represents the
CC glycosylated protein determination associated protein as described in
CC the invention.
XX
XX Sequence 585 AA;

Query Match 100.0%; Score 3103; DB 23; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.5e-255;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAFHFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
DB 1 DAHSEVAFHFKDLGEENFKALVLIAPAQYLOQCPEDHVKLVNEVTEFAKTCVADESAE 60
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDDNPNLPLVRPEV 120
DB 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNECFLOHKDDNPNLPLVRPEV 120
QY 121 DVMTAFHNEETFLKKLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLIP 180
DB 121 DVMTAFHNEETFLKKLYEIARRHPYFYAPPELLFAKRYKAAPTECCQAADKAACLIP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQFPKAEFAEVS KLIVTDLTK 240
DB 181 KLDELDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQFPKAEFAEVS KLIVTDLTK 240
QY 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMA 300
DB 241 VHTCECHGDLLECCADRDADLAKYICENQDSISSKLKECCPEKPLEKSHCIAEVENDEMA 300
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETTLK 360
QY 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCPELFEQGEYKFNALLVRYTKKVPQVST 420
DB 361 CAADPHECYAKVDFEFPKPLVEEPQNLKQNCPELFEQGEYKFNALLVRYTKKVPQVST 420
QY 421 PTLVEYSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480
DB 421 PTLVEYSRNLGKVGSKCKHPEAKMPCAEADYLSVVLNQLCVLHEKTPVSDRYTKCCTES 480
QY 481 LVNRRPCFSALEVDYETVYVPEKFAETTFPHADICTLSEKERQIKKQATALVELVKKPKAT 540

```
Db 481 LVNRRPCFSALEVDYVVKENATFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
QY 541 KEQLKAYMDDFAAFVEKCKADDKETCFABEGKLVAAQAALGL 585
Db 541 KEQLKAYMDDFAAFVEKCKADDKETCFABEGKLVAAQAALGL 585

RESULT 15
ABG63321
ID ABG63321 standard; protein; 585 AA.
AC ABG63321;
XX 27-AUG-2002 (first entry)
DE Human serum albumin (HSA) protein.
XX
KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
KW human serum albumin; HSA; cancer; reproductive disorder;
KW digestive disorder; immune disorder; endocrine disorder;
KW haematopoietic disorder; neural disorder; connective disorder;
KW cytostatic; antiinfertility; antiinflammatory; antitumor;
KW immunomodulator; anti-HIV; antidiabetic; haemostatic; nootropic;
KW neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
KW osteopathic; antiarthritic.
XX
OS Homo sapiens.
XX
XX WC200177137-AL.
XX
XX 18-OCT-2001.
XX
XX 12-APR-2001; 2001WO-US11988.
XX
XX 12-APR-2000; 2000US-229358P.
XX
XX 25-APR-2000; 2000US-199384P.
XX
XX 21-DEC-2000; 2000US-256931P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Haseltine WA;
XX
XX WPI; 2002-010886/01.
XX
XX N-PSDB; ABK93280.
XX
XX New fusion protein for treating disease e.g. diabetes comprises an
XX albumin fused to a therapeutic protein -
XX
XX Claim 1; Fig 15; 2102pp; English.
XX
XX The present invention relates to albumin fusion proteins comprising a
XX therapeutic protein X and human albumin (HA), also known as human serum
XX albumin, HSA). The proteins are useful for treating a disease or
XX disorder that may be modulated by therapeutic protein X. The albumin
XX extends the shelf-life of protein X, and may increase its biological
XX in vitro/in vivo activity. The protein is useful for treating and
XX diagnosing disorders such as cancer, reproductive disorders, digestive
XX disorders (e.g. Crohn's disease, ulcerative colitis), immune disorders
XX (e.g. acquired immunodeficiency syndrome, AIDS), endocrine disorders
XX (e.g. diabetes), haematopoietic disorders, neural disorders
XX (e.g. Alzheimer's, Parkinson's, Creutzfeldt-Jacob disease,
XX encephalomyelitis, meningitis, schizophrenia), and connective disorders
XX (e.g. osteoporosis, arthritis). The present sequence represents HSA
XX (HA) protein.
XX
XX Sequence 585 AA;
XX
XX Query Match 100.0%; Score 3103; DB 23; Length 585;
XX Best Local Similarity 100.0%; Pred. No. 9.5e-255;
XX Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 DAHSEVAHRFDLGEENKALVLAFAQLQCCPFEDHVKLVNEYTEFAKCVADESAE 60
```

```
Db 1 DAHSEVAHRFDLGEENKALVLAFAQLQCCPFEDHVKLVNEYTEFAKCVADESAE 60
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPNLPVLRPEV 120
Db 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBERNECFLOHKDDNPNLPVLRPEV 120
QY 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKYLIEIARRHPYFYAPELLFFAKRYKAAFTTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKORLKCSAQKFGERAFKAWAVARLSQRFPAEFAEYVKLVTDLTAK 240
Db 181 KLDELDEGKASSAKORLKCSAQKFGERAFKAWAVARLSQRFPAEFAEYVKLVTDLTAK 240
QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPILLEKSHCTAEVENDENPA 300
Db 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPILLEKSHCTAEVENDENPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYETARRHPDYSVILLRLAKTYETILEK 360
Db 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYETARRHPDYSVILLRLAKTYETILEK 360
QY 361 CAAADPHECTYAKVDFEKPFLVEEPQNLKQNCLEFQELGEYFQNALVRYTKVQVST 420
Db 361 CAAADPHECTYAKVDFEKPFLVEEPQNLKQNCLEFQELGEYFQNALVRYTKVQVST 420
QY 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEDYLSVILNOLCVLHEKTPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKHPKAKMPCAEDYLSVILNOLCVLHEKTPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYVVKENATFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDYVVKENATFTFHADICTLSEKERQIKQTALVELVHKPKAT 540
QY 541 KEQLKAYMDDFAAFVEKCKADDKETCFABEGKLVAAQAALGL 585
Db 541 KEQLKAYMDDFAAFVEKCKADDKETCFABEGKLVAAQAALGL 585
```

Search completed: August 31, 2003, 16:38:41
Job time : 88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 30 Seconds
(without alignments)
825.062 Million cell updates/sec

Title: US-09-833-041-18
Perfect score: 3103
Sequence: 1 DAKSEVAHFKDLGEENFK.....TCFAEKGKLVAAQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	585	1	US-08-153-799-14
2	3103	100.0	585	2	US-08-702-572-2
3	3103	100.0	585	3	US-08-769-746-2
4	3103	100.0	610	2	US-08-797-689-2
5	3103	100.0	783	1	US-08-256-938-2
6	3103	100.0	787	1	US-08-256-938-4
7	3103	100.0	787	2	US-08-797-689-16
8	3099	99.9	609	1	US-08-222-619-3
9	3099	99.9	609	1	US-08-433-037-4
10	3099	99.9	609	4	US-08-897-956A-2
11	3099	99.9	609	5	PCT-US95-04075-3
12	3095	99.7	978	4	US-08-897-956A-3
13	3093	99.7	585	1	US-08-448-196A-3
14	3093	99.7	585	2	US-08-984-176-1
15	2458.5	79.2	583	1	US-08-448-196A-5
16	2450.5	79.0	583	1	US-08-448-196A-6
17	2432.5	78.4	583	1	US-08-448-196A-7
18	2426	78.2	584	1	US-08-448-196A-7
19	2389	77.0	582	1	US-08-134-638-1
20	1249.5	40.3	609	1	US-08-222-619-4
21	1249.5	40.3	609	5	PCT-US95-04075-4
22	1206.5	38.9	590	2	US-08-377-309-2
23	1206.5	38.9	590	3	US-09-186-723-2
24	1206.5	38.9	590	4	US-08-505-012-5
25	1206.5	38.9	590	4	US-09-186-949A-3
26	1206.5	38.9	590	4	US-08-758-757-2
27	1206.5	38.9	590	5	PCT-US96-00996-5

Sequence 2, Appli
Sequence 8, Appli
Sequence 2, Appli
Sequence 24, Appli
Sequence 2, Appli
Sequence 7, Appli
Sequence 10, Appli
Sequence 8, Appli
Sequence 7, Appli
Sequence 10, Appli
Sequence 12, Appli
Sequence 8, Appli
Sequence 11, Appli
Sequence 9, Appli
Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-153-799-14
; Sequence 14, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Main Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8905916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Haln
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H332
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Region
; LOCATION: 369..419
; OTHER INFORMATION: /note= "Alternative C-termini of
; OTHER INFORMATION: HSA(1-n)"
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..585
; OTHER INFORMATION: /note= "Amino acid sequence of
; OTHER INFORMATION: natural HSA"
; US-08-153-799-14

Query Match      100.0%; Score 3103; DB 1; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLQCCPFEDHVKLVNVEVFCTCVADSSAE 60
Db 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLQCCPFEDHVKLVNVEVFCTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPLVRPEV 120
QY 121 DVMCTAFDNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 121 DVMCTAFDNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
QY 241 VHTCCGGDILLECADDRLADLAKYICENQDSISSKLKCECKPILLESKHCIAEVENDEMPA 300
Db 241 VHTCCGGDILLECADDRLADLAKYICENQDSISSKLKCECKPILLESKHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCNKYAEAKDVFLGMFLYETARRHPDYISVLLLLAKTYETTLTK 360
Db 301 DLPSLAADFVESKDVCNKYAEAKDVFLGMFLYETARRHPDYISVLLLLAKTYETTLTK 360
QY 361 CAADAPHECYAKVDFEKLVEEPONLTKQNCSELFQGLGEYKFQNALVRYTKKVPQYST 420
Db 361 CAADAPHECYAKVDFEKLVEEPONLTKQNCSELFQGLGEYKFQNALVRYTKKVPQYST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYLSVNLQCLVHLHETPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYLSVNLQCLVHLHETPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKQTAALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKQTAALVELVHKPKAT 540
QY 541 KEQLKAVMDDFAAFEKCKADDKETCTCFAEEGKKLVAASQAALGL 585
Db 541 KEQLKAVMDDFAAFEKCKADDKETCTCFAEEGKKLVAASQAALGL 585

RESULT 2
US-08-702-572-2
; Sequence 2, Application US/08702572
; Patent No. 5965386
; GENERAL INFORMATION:
; APPLICANT: Kerry-Williams, Sean M
; APPLICANT: Gilbert, Sarah C
; TITLE OF INVENTION: Yeast Strains and Modified Albumins
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Centeon L.L.C.
; STREET: 1020 First Avenue

```

```

; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-1310
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,572
; FILING DATE: 11-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 95/23857
; FILING DATE: 1-MAR-1995
; APPLICATION NUMBER: GB 9404270.2
; FILING DATE: 5-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Naomi Biswas
; REGISTRATION NUMBER: 38,384
; REFERENCE/DOCKET NUMBER: CE0114 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610/878/4294
; TELEFAX: 610/878/4221
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-702-572-2

```

```

Query Match      100.0%; Score 3103; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLQCCPFEDHVKLVNVEVFCTCVADSSAE 60
Db 1 DAHSEVAHFRKDLGEENFKALVLIATAFYAQYLQCCPFEDHVKLVNVEVFCTCVADSSAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPLVRPEV 120
Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPMLPLVRPEV 120
QY 121 DVMCTAFDNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 121 DVMCTAFDNEETFLKKYLYETARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
QY 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
Db 181 KLDELDEGKASSAKQRLKCSLQKFGGERAFKAWAVARLSQRPFAEFAVSKLVTDLTK 240
QY 241 VHTCCGGDILLECADDRLADLAKYICENQDSISSKLKCECKPILLESKHCIAEVENDEMPA 300
Db 241 VHTCCGGDILLECADDRLADLAKYICENQDSISSKLKCECKPILLESKHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCNKYAEAKDVFLGMFLYETARRHPDYISVLLLLAKTYETTLTK 360
Db 301 DLPSLAADFVESKDVCNKYAEAKDVFLGMFLYETARRHPDYISVLLLLAKTYETTLTK 360
QY 361 CAADAPHECYAKVDFEKLVEEPONLTKQNCSELFQGLGEYKFQNALVRYTKKVPQYST 420
Db 361 CAADAPHECYAKVDFEKLVEEPONLTKQNCSELFQGLGEYKFQNALVRYTKKVPQYST 420
QY 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYLSVNLQCLVHLHETPVSDRVTKCCTES 480
Db 421 PTLVEVSRNLGKVGSKCKKHPEAKMPCAEYLSVNLQCLVHLHETPVSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKQTAALVELVHKPKAT 540
Db 481 LVNRRPCFSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKQTAALVELVHKPKAT 540

```

QY 541 KEOLKAVMDFAAFVEKCKKADDETCFAEKGKLVAAQAALGL 585
Db 541 KEOLKAVMDFAAFVEKCKKADDETCFAEKGKLVAAQAALGL 585

RESULT 3

US-08-769-746-2
; Sequence 2, Application US/08769746
; Patent No. 6274305
; GENERAL INFORMATION:
; APPLICANT: Sonnschein, Carlos
; APPLICANT: Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/769,746
; FILING DATE: 19-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-769-746-2

Query Match 100.0%; Score 3103; DB 3; Length 585;
Best Local Similarity 100.0%; Pred. No. 9.3e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDLGKNEFKALVLIAPQYLQCCPPEDHVKLVNTEFAKTCVADDSAE 60
Db 1 DAHSEVAHFKDLGKNEFKALVLIAPQYLQCCPPEDHVKLVNTEFAKTCVADDSAE 60
QY 61 NCKDSLHTFLGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLRVEPV 120
Db 61 NCKDSLHTFLGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPNLPLRVEPV 120
QY 121 DVMCTAFHDNEETFLKKLYETARRHPYFYAPPELLFAKRYKAATECCQADKAACLLP 180
Db 121 DVMCTAFHDNEETFLKKLYETARRHPYFYAPPELLFAKRYKAATECCQADKAACLLP 180
QY 181 KLDELDRGKASSAKQRIKACSIQKFGERAFKANAVARLSQRPKAEFAEYSLKVIIDTK 240
Db 181 KLDELDRGKASSAKQRIKACSIQKFGERAFKANAVARLSQRPKAEFAEYSLKVIIDTK 240
QY 241 VTECHGDLLECDRADRIAKYICENQDSISSKKECCPKLLEKSHCIAFVENDMPA 300
Db 241 VTECHGDLLECDRADRIAKYICENQDSISSKKECCPKLLEKSHCIAFVENDMPA 300
QY 301 DLPSTAADFVSKDVCKNVAERKDVFLGMFLYEVARRHPDYSVLLLLRLAKYIETLEKC 360
Db 301 DLPSTAADFVSKDVCKNVAERKDVFLGMFLYEVARRHPDYSVLLLLRLAKYIETLEKC 360

QY 361 CAAADPHCYAKVDEFEKPLVEEPQNLKONCELFEOQLGEYKFNALLVRYTKKYPQVST 420
Db 361 CAAADPHCYAKVDEFEKPLVEEPQNLKONCELFEOQLGEYKFNALLVRYTKKYPQVST 420
QY 421 PTLVEVSRNLKGVSKCKKHPKAMPKADYLSVNLNQLCVLHEKTPYSDRVTKCCTES 480
Db 421 PTLVEVSRNLKGVSKCKKHPKAMPKADYLSVNLNQLCVLHEKTPYSDRVTKCCTES 480
QY 481 LVNRRPCFSALEVDYTVPKFEFNAETTFPHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db 481 LVNRRPCFSALEVDYTVPKFEFNAETTFPHADICTLSEKEROIKKOTALVELVKKPKAT 540
QY 541 KEOLKAVMDFAAFVEKCKKADDETCFAEKGKLVAAQAALGL 585
Db 541 KEOLKAVMDFAAFVEKCKKADDETCFAEKGKLVAAQAALGL 585

RESULT 4

US-08-797-689-2
; Sequence 2, Application US/08797689
; Patent No. 5876969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guillon, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3043
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-797-689-2

Query Match 100.0%; Score 3103; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 9.8e-287;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGEENFKALVLIAPAYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 60
Db 25 DAHSEVAHRFKDGEENFKALVLIAPAYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 84
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYIETARHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYIETARHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 264
QY 241 VTECCGDLLECADRADLAKYICENODSISKKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 265 VTECCGDLLECADRADLAKYICENODSISKKLKECCCKPILLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGFWFLYEAERHPDYSVLLRLAKTYETLEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGFWFLYEAERHPDYSVLLRLAKTYETLEKC 384
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 385 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 564
QY 541 KEQLKAVMDFFAAFEVKCKCKADDKTCFAEBGKKLVAASQAALGL 585
Db 565 KEQLKAVMDFFAAFEVKCKCKADDKTCFAEBGKKLVAASQAALGL 609

RESULT 5
US-08-256-938-2
; Sequence 2, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.0 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,938
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/01065
; FILING DATE: 31-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne

; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: ST92007-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3817
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 783 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-256-938-2

Query Match 100.0%; Score 3103; DB 1; Length 783;
Best Local Similarity 100.0%; Pred. No. 1.4e-286;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGEENFKALVLIAPAYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 60
Db 25 DAHSEVAHRFKDGEENFKALVLIAPAYLQOCPFFEDHVKLVNEVTEFAKTCVADSAE 84
QY 61 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 120
Db 85 NCDKSLHTLFGDKLCTVATIRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLVRPEV 144
QY 121 DVMCTAFHDNEETFLKKYIETARHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVMCTAFHDNEETFLKKYIETARHPYFYAPPELLFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 240
Db 205 KLDELDEGKASSAKQRLKASLOKFGERAFKAWAVARLSQRPKAEFAEYSKLVTDLT 264
QY 241 VTECCGDLLECADRADLAKYICENODSISKKLKECCCKPILLESKSHCIAEVENDEMPA 300
Db 265 VTECCGDLLECADRADLAKYICENODSISKKLKECCCKPILLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVLGFWFLYEAERHPDYSVLLRLAKTYETLEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVLGFWFLYEAERHPDYSVLLRLAKTYETLEKC 384
QY 361 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 420
Db 385 CAADPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNQALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPCAEDYLSVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDYTYVPKEFNAETFFHADICTLSEKEROIKKOTALVELVKKPKAT 564
QY 541 KEQLKAVMDFFAAFEVKCKCKADDKTCFAEBGKKLVAASQAALGL 585
Db 565 KEQLKAVMDFFAAFEVKCKCKADDKTCFAEBGKKLVAASQAALGL 609

RESULT 6
US-08-256-938-4
; Sequence 4, Application US/08256938
; Patent No. 5665863
; GENERAL INFORMATION:
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NEW POLYPEPTIDES HAVING GRANULOCYTE
; TITLE OF INVENTION: COLONY STIMULATING ACTIVITY, PREPARATION THEREOF AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.0 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/256,938
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/01065
FILING DATE: 31-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: ST92007-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3817
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-256-938-4

Query Match 100.0%; Score 3103; DB 1; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.4e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADSEAE 60
DB 203 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADSEAE 262
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNPLRVPEV 120
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNPLRVPEV 322
QY 121 DVMCTAFHDENEFTFLKKYIETARHPHYFYAPPELLFFAKRYKAATFCCQAADKAACLLP 180
DB 323 DVMCTAFHDENEFTFLKKYIETARHPHYFYAPPELLFFAKRYKAATFCCQAADKAACLLP 382
QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFAKAWARLSORPPKAEFAEVSPLYVDTLTK 240
DB 383 KLDELDEGKASSAKORLKCSLOKFGERAFAKAWARLSORPPKAEFAEVSPLYVDTLTK 442
QY 241 VTECHGDLLECAADDRADIACYICENQDSISSKLECCCKPILLESKSHCIAEVNDMPA 300
DB 443 VTECHGDLLECAADDRADIACYICENQDSISSKLECCCKPILLESKSHCIAEVNDMPA 502
QY 301 DLPSLAADFVSKDVCKNTAEAKDVLGMFLYETARHPDYSVVLLLRILAKYETTLTK 360
DB 503 DLPSLAADFVSKDVCKNTAEAKDVLGMFLYETARHPDYSVVLLLRILAKYETTLTK 562
QY 361 CAADPHCYAKVDEKPLVEPQNLKQNCFLFEOLGYKFNALIVRTTKVPOVST 420
DB 563 CAADPHCYAKVDEKPLVEPQNLKQNCFLFEOLGYKFNALIVRTTKVPOVST 622
QY 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 480
DB 623 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEYLSVLNQLCVLHEKTPVSDRVTKCTES 682
QY 481 LVNRPCTSALEVDVETVPKEFNAETFFHADICTLSEKROIKKQTALVELVKKPKAT 540
DB 683 LVNRPCTSALEVDVETVPKEFNAETFFHADICTLSEKROIKKQTALVELVKKPKAT 742
QY 541 KEOLKAVMDFAAEVFEKCKADDKETCFAEFGKLVNAASQAALGL 585
DB 743 KEOLKAVMDFAAEVFEKCKADDKETCFAEFGKLVNAASQAALGL 787

RESULT 7

US-08-797-689-16
Sequence 16, Application US/08797689
Patent No. 5876969
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
APPLICANT: Fournier, Alain
APPLICANT: Guittou, Jean-Dominique
APPLICANT: Jung, Gerard
APPLICANT: Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
TITLE OF INVENTION: CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESS: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA

ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/256,927

FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR93/00085

FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 787 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-797-689-16
Query Match 100.0%; Score 3103; DB 2; Length 787;
Best Local Similarity 100.0%; Pred. No. 1.4e-286;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADSEAE 60
DB 203 DAHSEVAHFKDGLGENFKALVLIAPQYLOQCPEFHVKLVNEVTEFAKTCVADSEAE 262
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNPLRVPEV 120
DB 263 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNPLRVPEV 322
QY 121 DVMCTAFHDENEFTFLKKYIETARHPHYFYAPPELLFFAKRYKAATFCCQAADKAACLLP 180
DB 323 DVMCTAFHDENEFTFLKKYIETARHPHYFYAPPELLFFAKRYKAATFCCQAADKAACLLP 382
QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFAKAWARLSORPPKAEFAEVSPLYVDTLTK 240
DB 383 KLDELDEGKASSAKORLKCSLOKFGERAFAKAWARLSORPPKAEFAEVSPLYVDTLTK 442

QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 443 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 502
QY 301 DLPSLAADFVESKDVCKNAEAKDVLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 360
DB 503 DLPSLAADFVESKDVCKNAEAKDVLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 562
QY 361 CAADAPHECYAKVDFEYKPLVEEPONLQKNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 563 CAADAPHECYAKVDFEYKPLVEEPONLQKNCLEFQOLGEYKFNALLVRYTKKVPQVST 622
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCCTES 480
DB 623 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCCTES 682
QY 481 LVNRRPCFSALEVDEYVYKPEFNAETFFHADICTLSEKEROIKKOTALVLYVKKPKAT 540
DB 683 LVNRRPCFSALEVDEYVYKPEFNAETFFHADICTLSEKEROIKKOTALVLYVKKPKAT 742
QY 541 KEOLKAVMDPFAAFVEKCKADDDKTCFAEGKKLVAASQAALGL 585
DB 743 KEOLKAVMDPFAAFVEKCKADDDKTCFAEGKKLVAASQAALGL 787

RESULT 8

US-08-222-619-3
; Sequence 3, Application US/08222619
; Patent No. 5652352
; GENERAL INFORMATION:
; APPLICANT: Lichenstein, Henri
; APPLICANT: Lyons, David
; APPLICANT: Wurfel, Mark
; APPLICANT: Wright, Samuel
; TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Center, Patent Operations/RRC
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: U.S.
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/222,619
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 609 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-222-619-3

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best Local Similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVARRKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEYFAKTCVADESAAE 60
DB 25 DAHSEVARRKDLGEENFKALVLIAPQYLQCCPFEDHVKLVNVEYFAKTCVADESAAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLRVREV 120

DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQPERNECFLOHKDDNPNLPLRVREV 144
QY 121 DVMCTAFHONETFTKKLYIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 180
DB 145 DVMCTAFHONETFTKKLYIEIARRHPYFYAPPELLFFAKRYKAAFTTECCQAADKAACLIP 204
QY 181 KLDELREDEKASAKORLKASLOKFGERAFKAWAVARLSORFFKAEFVSKLVTDLT 240
DB 205 KLDELREDEKASAKORLKASLOKFGERAFKAWAVARLSORFFKAEFVSKLVTDLT 264
QY 241 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNAEAKDVLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 360
DB 325 DLPSLAADFVESKDVCKNAEAKDVLGMFLYFYARRHPDYSVWLLRLAKTYETTTLEK 384
QY 361 CAADAPHECYAKVDFEYKPLVEEPONLQKNCLEFQOLGEYKFNALLVRYTKKVPQVST 420
DB 385 CAADAPHECYAKVDFEYKPLVEEPONLQKNCLEFQOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLQCVLHKTVPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDEYVYKPEFNAETFFHADICTLSEKEROIKKOTALVLYVKKPKAT 540
DB 505 LVNRRPCFSALEVDEYVYKPEFNAETFFHADICTLSEKEROIKKOTALVLYVKKPKAT 564
QY 541 KEOLKAVMDPFAAFVEKCKADDDKTCFAEGKKLVAASQAALGL 585
DB 565 KEOLKAVMDPFAAFVEKCKADDDKTCFAEGKKLVAASQAALGL 609

RESULT 9

US-08-433-037-4
; Sequence 4, Application US/08433037
; Patent No. 5707828
; GENERAL INFORMATION:
; APPLICANT: Sreekrishna, Kotikanyadan
; APPLICANT: Barr, Kathryn A.
; APPLICANT: Brietley, Russell A.
; APPLICANT: Thill, Gregory P.
; APPLICANT: Tschopp, Juerg F.
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN
; TITLE OF INVENTION: PICHIA PASTORIS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,037
; FILING DATE: 03-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9108Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-433-037-4

Query Match 99.9%; Score 3099; DB 1; Length 609;
Best local similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNVEYTFEAKTCVADESA 60
DB 25 DAHKSVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNVEYTFEAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNPNLRLVREPV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNPNLRLVREPV 144
QY 121 DVMTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELNDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFVSKLVDTLTK 240
DB 205 KLDELNDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFVSKLVDTLTK 264
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYIARRHPDYSVLLLRLLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYIARRHPDYSVLLLRLLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFLVPEEPQNLKQNCLEFQGEYKFQNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFLVPEEPQNLKQNCLEFQGEYKFQNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETTFHADICTLSEKEROIKQALVELVKKHPKAT 540
DB 505 LVNRRPCFSALEVDYTYVPKEFNAETTFHADICTLSEKEROIKQALVELVKKHPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 10
US-08-897-956A-2
Sequence 2, Application US/08897956A
Patent No. 6423512
GENERAL INFORMATION:
APPLICANT: Mary Ellen Digan
APPLICANT: Philip Lake
APPLICANT: Hermann Gram
TITLE OF INVENTION: Fusion Polypeptides
FILE REFERENCE: 600-7244/CPA
CURRENT APPLICATION NUMBER: US/08/897,956A
CURRENT FILING DATE: 1997-07-21
PRIOR APPLICATION NUMBER: 60/022,689
PRIOR FILING DATE: 1996-07-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 609
TYPE: PRT
ORGANISM: Homo Sapiens
US-08-897-956A-2

Query Match 99.9%; Score 3099; DB 4; Length 609;
Best local similarity 99.8%; Pred. No. 2.4e-286;
Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNVEYTFEAKTCVADESA 60
DB 25 DAHKSVAHRFKDLGEENFKALVLIAPQYLOQCFFEDHVKLVNVEYTFEAKTCVADESA 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNPNLRLVREPV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPNERNECFLOHKDDNPNLRLVREPV 144
QY 121 DVMTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 145 DVMTAFHDNEETFLKKYLIEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELNDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFVSKLVDTLTK 240
DB 205 KLDELNDEGKASSAKQRLKCSLQKFGERAFKAWAVARLSQRPFAEFVSKLVDTLTK 264
QY 241 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTECCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYIARRHPDYSVLLLRLLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVFGLMFLEYIARRHPDYSVLLLRLLAKTYETTLK 384
QY 361 CAAADPHECYAKVDFEFLVPEEPQNLKQNCLEFQGEYKFQNALLVRYTKVPQVST 420
DB 385 CAAADPHECYAKVDFEFLVPEEPQNLKQNCLEFQGEYKFQNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYTYVPKEFNAETTFHADICTLSEKEROIKQALVELVKKHPKAT 540
DB 505 LVNRRPCFSALEVDYTYVPKEFNAETTFHADICTLSEKEROIKQALVELVKKHPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 585
DB 565 KEQLKAVMDDDFAAFVEKCKCKADDDKTCFAEKGKLVAAASQAALGL 609

RESULT 11
PCT-US95-04075-3
Sequence 3, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 609 amino acids

; TYPE: amino acid
 ; STRANDEDNESS: unknown
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 PCT-US95-04075-3

Query Match 99.9%; Score 3099; DB 5; Length 609;
 Best Local Similarity 99.8%; Pred. No. 2.4e-286;
 Matches 584; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 DAHSEVAHFRKDLGEENFKALVLIATAQYLOQCPEDHVKLVNVEFVAKTCVADESAE 60
 25 DAHSEVAHFRKDLGEENFKALVLIATAQYLOQCPEDHVKLVNVEFVAKTCVADESAE 84
 61 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120
 85 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 144
 121 DVMTAFPHDNEETFLKKLYIEIARRHPFYFAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
 145 DVMTAFPHDNEETFLKKLYIEIARRHPFYFAPPELLFAKRYKAAFTCCQAADKAAACLLP 204
 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLYK 240
 205 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLYK 264
 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
 265 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 324
 301 DLPSLAADRVESKDVCKNTAEAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTTLEK 360
 325 DLPSLAADRVESKDVCKNTAEAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTTLEK 384
 361 CAAADPHECYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 420
 385 CAAADPHECYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 444
 421 PTLVEVSRLNGKVGSKCKKHPEAKMPCAEEDYLSVNLQCVLHKTTPVSDRVTKCCTES 480
 445 PTLVEVSRLNGKVGSKCKKHPEAKMPCAEEDYLSVNLQCVLHKTTPVSDRVTKCCTES 504
 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 564
 541 KEOLKAYMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALGL 585
 565 KEOLKAYMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALGL 609

RESULT 12
 US-08-897-956A-3
 ; Sequence 3, Application US/08897956A
 ; Patent No. 6423512
 ; GENERAL INFORMATION:
 ; APPLICANT: Mary Ellen Digan
 ; APPLICANT: Philip Lake
 ; APPLICANT: Hermann Gram
 ; TITLE OF INVENTION: Fusion Polypeptides
 ; FILE REFERENCE: 600-7244/CFA
 ; CURRENT APPLICATION NUMBER: US/08/897,956A
 ; CURRENT FILING DATE: 1997-07-21
 ; PRIOR APPLICATION NUMBER: 60/022,689
 ; PRIOR FILING DATE: 1996-07-26
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 978
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Fusion polypeptide

US-08-897-956A-3

Query Match 99.7%; Score 3095; DB 4; Length 978;
 Best Local Similarity 99.8%; Pred. No. 1.1e-285;
 Matches 583; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 DAHSEVAHFRKDLGEENFKALVLIATAQYLOQCPEDHVKLVNVEFVAKTCVADESAE 60
 212 DAHSEVAHFRKDLGEENFKALVLIATAQYLOQCPEDHVKLVNVEFVAKTCVADESAE 271
 61 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 120
 272 NCKSLHTLFGDKLCIVATLRETYGEMADCCAKQEPERNECFLOHKDNPMLPLVRPEV 331
 121 DVMTAFPHDNEETFLKKLYIEIARRHPFYFAPPELLFAKRYKAAFTCCQAADKAAACLLP 180
 332 DVMTAFPHDNEETFLKKLYIEIARRHPFYFAPPELLFAKRYKAAFTCCQAADKAAACLLP 391
 181 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLYK 240
 392 KLDELDEGKASSAKORLKCASLQKFGERAFKAWAVARLSQRPKAEFAEYVKLVITDLYK 451
 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 300
 452 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECEKPLLEKSHCIAEVENDEMPA 511
 301 DLPSLAADRVESKDVCKNTAEAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTTLEK 360
 512 DLPSLAADRVESKDVCKNTAEAKDVFGLMFLYETARRHPDYSVLLRLAKTYETTTLEK 571
 361 CAAADPHECYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 420
 572 CAAADPHECYAKVDFEDEFKPLVEEPONLIKONCELFQOLGEYFQNALVRYTKKVPQVST 631
 421 PTLVEVSRLNGKVGSKCKKHPEAKMPCAEEDYLSVNLQCVLHKTTPVSDRVTKCCTES 480
 632 PTLVEVSRLNGKVGSKCKKHPEAKMPCAEEDYLSVNLQCVLHKTTPVSDRVTKCCTES 691
 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
 692 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 751
 541 KEOLKAYMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALG 584
 752 KEOLKAYMDDFAAFVEKCKCADDKTCFAEBGKKLVAASQAALG 795

RESULT 13
 US-08-448-196A-3
 ; Sequence 3, Application US/08448196A
 ; Patent No. 5780594
 ; GENERAL INFORMATION:
 ; APPLICANT: CARTER, DANIEL C.
 ; TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
 ; TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
 ; TITLE OF INVENTION: RELATED PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NASA
 ; STREET: MARSHALL SPACE FLIGHT CENTER
 ; CITY: HUNTSVILLE
 ; STATE: ALABAMA
 ; COUNTRY: USA
 ; ZIP: 35812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/448,196A
 ; FILING DATE: 23-MAY-1995
 ; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: BROAD JR., ROBERT L.
 REGISTRATION NUMBER: 18,757
 REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 205-544-0021
 TELEFAX: 205-544-0258
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 585 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 US-08-448-196A-3

Query Match 99.7%; Score 3093; DB 1; Length 585;
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLETTYGENADCCAKQBPENECFLOHKDDNPNLRLVPRPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLETTYGENADCCAKQBPENECFLOHKDDNPNLRLVPRPV 120

QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
 DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLIP 180

QY 181 KLDELDEGKASSAKQRLKCSAQKGFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKQRLKCSAQKGFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240

QY 241 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVNDEMFA 300
 DB 241 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVNDEMFA 300

QY 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLRLAKTYETTLK 360
 DB 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLRLAKTYETTLK 360

QY 361 CAADHPHCYAKVDFEFPPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADHPHCYAKVDFEFPPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEYSRNLGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
 DB 421 PTLVEYSRNLGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480

QY 481 LVNRRPCFSALEVDYETYPKEFNAETFFHADICTLSEKEROIKQKOTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALEVDYETYPKEFNAETFFHADICTLSEKEROIKQKOTALVELVKKPKAT 540

QY 541 KEQLKAVMDPFAAFVEKCKCKADDETCFAEKGKLVASQAALGL 585
 DB 541 KEQLKAVMDPFAAFVEKCKCKADDETCFAEKGKLVASQAALGL 585

RESULT 14

US-08-984-176-1

Sequence 1, Application US/08984176
 Patent No. 5948609

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C
 APPLICANT: HO, JOSEPH X
 APPLICANT: RUKER, FLORIAN
 TITLE OF INVENTION: OXYGEN-TRANSPORTING ALBUMIN-BASED BLOOD REPLACEMENT
 TITLE OF INVENTION: COMPOSITION AND BLOOD VOLUME EXPANDER

FILE REFERENCE: 08/984,176
 CURRENT APPLICATION NUMBER: US/08/984,176
 CURRENT FILING DATE: 1997-12-03

NUMBER OF SEQ ID NOS: 1
 SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 585

TYPE: PRT

ORGANISM: Homo sapiens

US-08-984-176-1

Query Match 99.7%; Score 3093; DB 2; Length 585;
 Best Local Similarity 99.7%; Pred. No. 8.3e-286;
 Matches 583; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60
 DB 1 DAHSEVAHRFKDLGEENFKALVLIAPAOYLQCCPFEDHVKLVNEVTEFAKTCVADESAE 60

QY 61 NCDKSLHTLFGDKLCTVATLETTYGENADCCAKQBPENECFLOHKDDNPNLRLVPRPV 120
 DB 61 NCDKSLHTLFGDKLCTVATLETTYGENADCCAKQBPENECFLOHKDDNPNLRLVPRPV 120

QY 121 DVMTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
 DB 121 DVMTAFHDNEETFLKYLVEIARRHPYFVAPPELLFFAKRYKAAFTCCQAADKAACLIP 180

QY 181 KLDELDEGKASSAKQRLKCSAQKGFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240
 DB 181 KLDELDEGKASSAKQRLKCSAQKGFGERAFKAWAVARLSORFPKAEFAVSKLVTDLTK 240

QY 241 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVNDEMFA 300
 DB 241 VHTCCGDLLECCADRADLAKYICENODSISSKLKECCPKLLEKSHCIAEVNDEMFA 300

QY 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLRLAKTYETTLK 360
 DB 301 DLPFLAADVESEKDYCKNYAEAKDVFLGMFLYEYARRHPDYSVWLLRLAKTYETTLK 360

QY 361 CAADHPHCYAKVDFEFPPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
 DB 361 CAADHPHCYAKVDFEFPPLVEEPQNLKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420

QY 421 PTLVEYSRNLGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
 DB 421 PTLVEYSRNLGVSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480

QY 481 LVNRRPCFSALEVDYETYPKEFNAETFFHADICTLSEKEROIKQKOTALVELVKKPKAT 540
 DB 481 LVNRRPCFSALEVDYETYPKEFNAETFFHADICTLSEKEROIKQKOTALVELVKKPKAT 540

QY 541 KEQLKAVMDPFAAFVEKCKCKADDETCFAEKGKLVASQAALGL 585
 DB 541 KEQLKAVMDPFAAFVEKCKCKADDETCFAEKGKLVASQAALGL 585

RESULT 15

US-08-448-196A-5

Sequence 5, Application US/08448196A

Patent No. 5780594

GENERAL INFORMATION:

APPLICANT: CARTER, DANIEL C.
 TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN FRAGMENTS
 TITLE OF INVENTION: CONTAINING SPECIFIC BINDING REGIONS OF SERUM ALBUMIN OR
 TITLE OF INVENTION: RELATED PROTEINS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NASA
 STREET: MARSHALL SPACE FLIGHT CENTER
 CITY: HUNTSVILLE
 STATE: ALABAMA
 COUNTRY: USA
 ZIP: 35812

Job time : 32 secs

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/448,196A
FILING DATE: 23-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: BROAD JR., ROBERT L.
REGISTRATION NUMBER: 18,757
REFERENCE/DOCKET NUMBER: XX/MFS-28402-2
TELEPHONE: 205-544-0021
TELEFAX: 205-544-0258
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-448-196A-5

Query Match 79.2%; Score 2458.5; DB 1; Length 583;
Best local similarity 75.8%; Pred. No. 2.1e-225;
Matches 442; Conservative 70; Mismatches 70; Indels 1; Gaps 1;
QY 1 DAHKEVAREKDLGEENFKALVLIARAFYLOQCPFFEDHVKLVNEVTEFAKTCVADSEAE 60
Db 1 DTHKSEIARHNDGEKHFGLVAFSOTLOQCPFFEDHVKLVNEVTEFAKTCVADSEAE 60
QY 61 NCDKSLHTLFGDKLTVALRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEV 120
Db 61 NCDKSLHTLFGDKLTVALRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPEV 120
QY 121 DVNCFADHNBETFLKYLVEIARRHPYIAPELLFAKRYKAAFTCCQAADKACLLP 180
Db 120 DAQCAAFQEDPKFLGYLVEIARRHPYIAPELLFAKRYKAAFTCCQAADKACLLP 180
QY 181 KLDELDFEGKASSAKQRLKASLOKTERAFKAVARLSORPPKAEFAEVSCLVDTLTK 240
Db 180 KLDAKERILLSSAKERLKCSSQNGERAVKANSVALSQPKADPAEVSCLVDTLTK 240
QY 241 VHFECCHGLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDENMPA 300
Db 240 VHFECCHGLLECADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVNDENMPA 300
QY 301 DLPSLAADPVESKQCKNYAEKDVFLGMFLVEYARRHPDYSVVLLIRLAKTYETTLK 360
Db 300 DIPALAADPVESKQCKNYAEKDVFLGMFLVEYARRHPDYSVVLLIRLAKTYETTLK 360
QY 361 CAAADPHECYARVDEEKPVEPPQNLKQNCLEFQELGEYKFNALLVRYTKKVPQVST 420
Db 360 CAAADPACRYTVEDQFTPLVEBPKSLVKNCDLFEVEGYDFQNALIVRYTKKAPQVST 420
QY 421 PTLVEVSRNLGVSKCKHPEAKRMPDCAEDYLSVLNQLCVLHEKTPVSDRYTKCTES 480
Db 420 PTLVEIGRTLGVSRCKLPSESLPCSENHALALNRLCVLHEKTPVSEKITRCCWDS 479
QY 481 LVNRRCFSALREVDYTPVPEFNAETFTFHADICTLSEKROIKKOTALVELVKKPKAT 540
Db 480 LAERRCFSALREVDYTPVPEFNAETFTFHADICTLSEKROIKKOTALVELVKKPKAT 540
QY 541 KEQLKAVMDFAAFVEKCKADDKETCFAEKCKLVAASQAAL 583
Db 540 KEQLKAVMDFAAFVEKCKADDKETCFAEKCKLVAASQAAL 583

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:53 ; Search time 28 Seconds
(without alignments)
2857.764 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103
Sequence: 1 DAKSEVAHRFKOLGENFK.....TCFAEKGKLVAAASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 510680 seqs, 136781880 residues

Total number of hits satisfying chosen parameters: 510680

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

```
1: /cgn2.6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2.6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2.6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2.6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2.6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2.6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2.6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2.6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2.6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2.6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2.6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2.6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2.6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2.6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2.6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2.6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2.6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2.6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	3103	100.0	585	10	US-09-929-552-2	Sequence 2, Appli
2	3103	100.0	585	11	US-09-932-613-445	Sequence 445, App
3	3103	100.0	585	11	US-09-984-010-26	Sequence 26, Appl
4	3103	100.0	585	11	US-09-833-041-18	Sequence 18, Appl
5	3103	100.0	585	12	US-10-153-604A-5	Sequence 5, Appli
6	3103	100.0	585	14	US-10-153-064-5	Sequence 2, Appli
7	3103	100.0	604	11	US-09-984-010-7	Sequence 7, Appli
8	3103	100.0	609	11	US-09-919-039-370	Sequence 370, App
9	3103	100.0	609	12	US-10-153-604A-7	Sequence 7, Appli
10	3103	100.0	609	14	US-10-153-064-7	Sequence 7, Appli
11	3103	100.0	610	15	US-09-984-186-2	Sequence 2, Appli
12	3103	100.0	610	15	US-10-237-667-2	Sequence 2, Appli
13	3103	100.0	610	15	US-10-237-708-2	Sequence 2, Appli
14	3103	100.0	610	15	US-10-237-866-2	Sequence 2, Appli
15	3103	100.0	610	15	US-10-237-871-2	Sequence 2, Appli

```
16 3103 100.0 610 15 US-10-237-624-2 Sequence 2, Appli
17 3103 100.0 651 12 US-10-153-604A-133 Sequence 133, App
18 3103 100.0 651 14 US-10-153-064-133 Sequence 133, App
19 3103 100.0 652 12 US-10-153-604A-132 Sequence 132, App
20 3103 100.0 652 14 US-10-153-064-132 Sequence 132, App
21 3103 100.0 653 12 US-10-153-604A-131 Sequence 131, App
22 3103 100.0 653 14 US-10-153-064-131 Sequence 131, App
23 3103 100.0 656 12 US-10-153-604A-130 Sequence 130, App
24 3103 100.0 656 14 US-10-153-064-130 Sequence 130, App
25 3103 100.0 676 12 US-10-153-604A-127 Sequence 127, App
26 3103 100.0 676 14 US-10-153-604A-127 Sequence 127, App
27 3103 100.0 676 14 US-10-153-064-127 Sequence 127, App
28 3103 100.0 677 12 US-10-153-604A-125 Sequence 125, App
29 3103 100.0 677 14 US-10-153-604A-125 Sequence 125, App
30 3103 100.0 677 14 US-10-153-064-125 Sequence 125, App
31 3103 100.0 680 12 US-10-153-604A-123 Sequence 123, App
32 3103 100.0 680 14 US-10-153-064-123 Sequence 123, App
33 3103 100.0 787 10 US-09-984-186-16 Sequence 16, Appl
34 3103 100.0 787 15 US-10-237-667-16 Sequence 16, Appl
35 3103 100.0 787 15 US-10-237-708-16 Sequence 16, Appl
36 3103 100.0 787 15 US-10-237-866-16 Sequence 16, Appl
37 3103 100.0 787 15 US-10-237-871-16 Sequence 16, Appl
38 3103 100.0 787 15 US-10-237-624-16 Sequence 16, Appl
39 3103 100.0 788 15 US-10-073-118-26 Sequence 26, Appl
40 3092.5 99.7 652 12 US-10-153-604A-96 Sequence 96, Appl
41 3092.5 99.7 652 12 US-10-153-604A-99 Sequence 99, Appl
42 3092.5 99.7 652 12 US-10-153-604A-105 Sequence 105, App
43 3092.5 99.7 652 14 US-10-153-064-96 Sequence 96, Appl
44 3092.5 99.7 652 14 US-10-153-064-99 Sequence 99, Appl
45 3092.5 99.7 652 14 US-10-153-064-105 Sequence 105, App
```

ALIGNMENTS

```
RESULT 1
US-09-929-552-2
; Sequence 2, Application US/09929552
; Patent No. US20020123080A1
; GENERAL INFORMATION:
; APPLICANT: Sonnschein, Carlos
; Soto, Ana M.
; TITLE OF INVENTION: Inhibiting Proliferation of Cancer Cells
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/929,552
; FILING DATE: 14-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/769,746
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: MBRI-02584
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
```

; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-929-552-2

Query Match 100.0%; Score 3103; DB 10; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-270;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOQCPPEDEHVKLVNEVTEFAKTCVADESAE 60
 Db 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOQCPPEDEHVKLVNEVTEFAKTCVADESAE 60
 QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
 Db 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
 QY 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAATTECCQAADKAACLLP 180
 Db 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAATTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
 Db 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVLLLLRLAKTYETTLK 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVLLLLRLAKTYETTLK 360
 QY 361 CAADAPHECYAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420
 Db 361 CAADAPHECYAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLNGKVGSKCKHPKAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480
 Db 421 PTLVEVSRLNGKVGSKCKHPKAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKHPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKHPKAT 540
 QY 541 KEOLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAQAAALGL 585
 Db 541 KEOLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAQAAALGL 585

RESULT 2
 US-09-932-613-445
 ; Sequence 445, Application US/09932613
 ; Publication No. US20030091565A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Human Genome Sciences, Inc.
 ; APPLICANT: Belter, James P.
 ; APPLICANT: Potter, M. Daniel
 ; APPLICANT: Fleming, Tony J.
 ; APPLICANT: Rosen, Craig A.
 ; TITLE OF INVENTION: BINDING POLYPEPTIDES AND METHODS BASED THEREON
 ; FILE REFERENCE: DXX-025.1 PCT; DXX-025.1 US
 ; CURRENT APPLICATION NUMBER: US/09/932/613
 ; CURRENT FILING DATE: 2001-08-17
 ; NUMBER OF SEQ ID NOS: 458
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 445
 ; LENGTH: 585
 ; TYPE: PRT
 ; ORGANISM: HomoSapiens
 US-09-932-613-445

Query Match 100.0%; Score 3103; DB 11; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-270;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOQCPPEDEHVKLVNEVTEFAKTCVADESAE 60
 Db 1 DAHSEVAHFRKDLGEENFKALVLIAPAQYLOQCPPEDEHVKLVNEVTEFAKTCVADESAE 60
 QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
 Db 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRLVPEV 120
 QY 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAATTECCQAADKAACLLP 180
 Db 121 DVMCTAPHDNEETFLKKLYIEIARRHPYFYAPPELLFFAKRYKAATTECCQAADKAACLLP 180
 QY 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
 Db 181 KLDELDEGKASSAKORLKCSLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
 QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 Db 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
 QY 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVLLLLRLAKTYETTLK 360
 Db 301 DLPSLAADFVESKDVCKNYAEAKOVFLGMFLYIYARRHPDYSVLLLLRLAKTYETTLK 360
 QY 361 CAADAPHECYAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420
 Db 361 CAADAPHECYAKVDFEFLPEEONLIKONCELFOLGEYKFNALLVRYTKKVPQVST 420
 QY 421 PTLVEVSRLNGKVGSKCKHPKAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480
 Db 421 PTLVEVSRLNGKVGSKCKHPKAKMPCAEEDYLSVVLNQLCVLHEKTPVSDRVTKCOTES 480
 QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKHPKAT 540
 Db 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKHPKAT 540
 QY 541 KEOLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAQAAALGL 585
 Db 541 KEOLKAVMDFFAAFEVKCKCKADDDKTCFAEKGKLVAAQAAALGL 585

RESULT 3
 US-09-984-010-26
 ; Sequence 26, Application US/09984010
 ; Publication No. US20030104578A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ballance, David James
 ; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 ; AND SERUM ALBUMIN
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP
 ; STREET: 1300 I Street, NW
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/984.010
 ; FILING DATE: 21-May-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/091,873
 ; FILING DATE: 25-JUN-1998
 ; APPLICATION NUMBER: PCT/GB96/03164

/ FILING DATE: 19-DEC-1996
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 585 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: <unknown>
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKNVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKNVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
QY 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVILLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVILLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKPVOST 420
DB 361 CAADAPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKPVOST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAVFEKCKCKADKCTCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAVFEKCKCKADKCTCFABEGKKLVAASQAALGL 585

RESULT 4
US-09-833-041-18
/ Sequence 18, Application US/09833041
/ Publication No. US20030125247A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen, Craig A.
/ APPLICANT: Haseltine, William A.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PF545
/ CURRENT APPLICATION NUMBER: US/09/833,041
/ PRIOR FILING DATE: 2001-04-12
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-984-010-26

/ PRIOR APPLICATION NUMBER: 60/199,384

/ PRIOR FILING DATE: 2000-04-25

/ NUMBER OF SEQ ID NOS: 79

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 18

/ LENGTH: 585

/ TYPE: PRT

/ ORGANISM: Homo Sapiens

/ US-09-833-041-18

Query Match 100.0%; Score 3103; DB 11; Length 585;
Best Local Similarity 100.0%; Pred. No. 1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKNVNEVTEFAKTCVADESAAE 60
DB 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEFHVKNVNEVTEFAKTCVADESAAE 60
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
DB 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVRPEV 120
QY 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
DB 121 DVMTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLIP 180
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
DB 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSQRFPAEFAEVS KLVTDLTK 240
QY 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
DB 241 VHTCCGHDLLLECADRADLAKYICENQDSISSKLEKCEKPELLEKSHCIAEVENDEMPA 300
QY 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVILLRLAKTYETTLK 360
DB 301 DLPSLAADFVESKDVCKNVAEAKDVFLGMFLYEYARRHPDYSVILLRLAKTYETTLK 360
QY 361 CAADAPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKPVOST 420
DB 361 CAADAPHECYAKVDFEFKPLVEEPQNLKQNCLEFQOLGEYKFNALLVRYTKKPVOST 420
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
DB 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRYTKCTES 480
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540
DB 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKQTALVELVKKPKAT 540
QY 541 KEQLKAVMDDFAAVFEKCKCKADKCTCFABEGKKLVAASQAALGL 585
DB 541 KEQLKAVMDDFAAVFEKCKCKADKCTCFABEGKKLVAASQAALGL 585

RESULT 5

US-10-153-604A-5

/ Sequence 5, Application US/10153604A

/ Publication No. US20030143191A1

/ GENERAL INFORMATION:

/ APPLICANT: Bell et al.

/ TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins

/ FILE REFERENCE: PF556

/ CURRENT APPLICATION NUMBER: US/10/153,604A

/ PRIOR FILING DATE: 2002-05-24

/ PRIOR APPLICATION NUMBER: 60/293,212

/ NUMBER OF SEQ ID NOS: 137

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 5

/ LENGTH: 585

/ TYPE: PRT

/ ORGANISM: Homo Sapiens

US-10-153-604A-5

Query Match 100.0%; Score 3103; DB 12; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-270;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLSGNEFKALVLIATAFYAQYLQCCPPEDHVKLVNEVTEFAKTCVADSAE 60
 |||||
 Db 1 DAHSEVAHRFKDLSGNEFKALVLIATAFYAQYLQCCPPEDHVKLVNEVTEFAKTCVADSAE 60
 |||||
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLRVPEV 120
 |||||
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLRVPEV 120
 |||||
 QY 121 DVMCTAFHDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 |||||
 Db 121 DVMCTAFHDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 |||||
 QY 181 KIDELRDEGKASSAKORLKCASLOKFGGRAPKAWARLSQRPFAEFAEVS KLVTDLTK 240
 |||||
 Db 181 KIDELRDEGKASSAKORLKCASLOKFGGRAPKAWARLSQRPFAEFAEVS KLVTDLTK 240
 |||||
 QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 |||||
 Db 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 |||||
 QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYSVLLRLAKTYETTLTK 360
 |||||
 Db 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYSVLLRLAKTYETTLTK 360
 |||||
 QY 361 CAADPHCEYAKYDFDEKPLVEEPONLTKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
 |||||
 Db 361 CAADPHCEYAKYDFDEKPLVEEPONLTKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
 |||||
 QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480
 |||||
 Db 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480
 |||||
 QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
 |||||
 Db 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
 |||||
 QY 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585
 |||||
 Db 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585
 |||||

RESULT 6

US-10-153-064-5
 ; Sequence 5, Application US/10153064
 ; Publication No. US20020142814A1

GENERAL INFORMATION:
 ; APPLICANT: Bell et al.

; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
 ; FILE REFERENCE: PF556

; CURRENT APPLICATION NUMBER: US/10/153,064
 ; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: 60/293,212
 ; PRIOR FILING DATE: 2001-05-25

; NUMBER OF SEQ ID NOS: 137
 ; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5

; LENGTH: 585

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-10-153-064-5

Query Match 100.0%; Score 3103; DB 14; Length 585;
 Best Local Similarity 100.0%; Pred. No. 1e-270;
 Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDLSGNEFKALVLIATAFYAQYLQCCPPEDHVKLVNEVTEFAKTCVADSAE 60
 |||||

Db 1 DAHSEVAHRFKDLSGNEFKALVLIATAFYAQYLQCCPPEDHVKLVNEVTEFAKTCVADSAE 60
 |||||
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLRVPEV 120
 |||||
 Db 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPERNECFLOHKDDNPMLPLRVPEV 120
 |||||
 QY 121 DVMCTAFHDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 |||||
 Db 121 DVMCTAFHDNEETFLKKYLYEIARHPYFYAPPELLFFAKRYKAAFTCCQAADKAACLLP 180
 |||||
 QY 181 KIDELRDEGKASSAKORLKCASLOKFGGRAPKAWARLSQRPFAEFAEVS KLVTDLTK 240
 |||||
 Db 181 KIDELRDEGKASSAKORLKCASLOKFGGRAPKAWARLSQRPFAEFAEVS KLVTDLTK 240
 |||||
 QY 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 |||||
 Db 241 VHTCCCHGDLLECADRADLAKYICENODSISSKLKECCCKPILKSKSHCIAEVENDEMPA 300
 |||||
 QY 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYSVLLRLAKTYETTLTK 360
 |||||
 Db 301 DLPSLAADFVESKDVCKNTAEAKDVFLGMFLYFYARHPDYSVLLRLAKTYETTLTK 360
 |||||
 QY 361 CAADPHCEYAKYDFDEKPLVEEPONLTKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
 |||||
 Db 361 CAADPHCEYAKYDFDEKPLVEEPONLTKONCELFQOLGEYKFNALLVRYTKKVPQVST 420
 |||||
 QY 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480
 |||||
 Db 421 PTLVEVSRNLGKVGSKCKHPKAPCAEDYLSVNLNOLCVLHKTVPVSDRVTKCCTES 480
 |||||
 QY 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
 |||||
 Db 481 LVNRRPCFSALEVDYTYVPKEFNAETFTFHADICTLSEKEROIKKOTALVELVHKPKAT 540
 |||||
 QY 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585
 |||||
 Db 541 KEOLKAVMDFAAFVEKCKKADKTCFAEGBGKKLVAASQAALGL 585
 |||||

RESULT 7

US-09-984-010-7
 ; Sequence 7, Application US/09984010
 ; Publication No. US20030104578A1

GENERAL INFORMATION:
 ; APPLICANT: Ballance, David James

; TITLE OF INVENTION: RECOMBINANT FUSION PROTEINS TO GROWTH HORMONE
 ; AND SERUM ALBUMIN

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESS: FINNEGAN, HENDERSON, FARABOW, GARRETT & DUNNER, LLP

STREET: 1300 I Street, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/984,010

FILING DATE: 21-May-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/091,873

FILING DATE: 25-JUN-1998

APPLICATION NUMBER: PCT/GB96/03164

FILING DATE: 19-DEC-1996

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 604 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

```

;
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-984-010-7

Query Match          100.0%; Score 3103; DB 11; Length 604;
Best Local Similarity 100.0%; Pred. No. 1.le-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGEENFKALVLIAPAFQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 60
DB 20 DAHKSVAHRFKDGLGEENFKALVLIAPAFQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 79
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
DB 80 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 139
QY 121 DVNCTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 140 DVNCTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 199
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 240
DB 200 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 259
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 260 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 319
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETLEKC 360
DB 320 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETLEKC 379
QY 361 CAADAPHECYAKVDFEPLVEEPQNLIKONCELFGEQKVFONALLVRYTKKVPQVST 420
DB 380 CAADAPHECYAKVDFEPLVEEPQNLIKONCELFGEQKVFONALLVRYTKKVPQVST 439
QY 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 440 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 499
QY 481 LVNRRPCFSALVEVDYTYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKKPKAT 540
DB 500 LVNRRPCFSALVEVDYTYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKKPKAT 559
QY 541 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 585
DB 560 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 604

RESULT 8
US-09-919-039-370
; Sequence 370, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 370
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 088957CD1
```

```

US-09-919-039-370

Query Match          100.0%; Score 3103; DB 11; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.le-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGEENFKALVLIAPAFQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 60
DB 25 DAHKSVAHRFKDGLGEENFKALVLIAPAFQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 84
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 120
DB 85 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQBPNERNECFLOHKDDNPNLRLVRPEV 144
QY 121 DVNCTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
DB 145 DVNCTAFHDNEETFLKKYLXIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKQRLKCSLQKFGERAFAKAWAVARLSORFPKAEFVSKLVTDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETLEKC 360
DB 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYSVLLLRKAKTYETLEKC 384
QY 361 CAADAPHECYAKVDFEPLVEEPQNLIKONCELFGEQKVFONALLVRYTKKVPQVST 420
DB 385 CAADAPHECYAKVDFEPLVEEPQNLIKONCELFGEQKVFONALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRNLKGVSKCKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALVEVDYTYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKKPKAT 540
DB 505 LVNRRPCFSALVEVDYTYVPKEFNAETTFHADICTLSEKERQIKKQATALVELVKKPKAT 564
QY 541 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDDFAAFEVKCKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 9
US-10-153-604A-7
; Sequence 7, Application US/10153604A
; Publication No. US20030143191A1
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,604A
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-604A-7

Query Match          100.0%; Score 3103; DB 12; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.le-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHKSVAHRFKDGLGEENFKALVLIAPAFQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 60
DB 11 DAHKSVAHRFKDGLGEENFKALVLIAPAFQYLOQCFFEDHVKLVNEVTEFAKTCVADESA 60
```

```
Db 25 DAHSEVAHRFKDGLGEENFKALVLIAPAFQYLQCCPPEDHVLVNEVTEFAKTCVADSAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPLRVPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPLRVPEV 144
QY 121 DVMTAFPHDNEETFLKKLYIETARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 180
Db 145 DVMTAFPHDNEETFLKKLYIETARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 204
QY 181 KLDELDEGKASAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVS KLVDLTJK 240
Db 205 KLDELDEGKASAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVS KLVDLTJK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYETARRHPDYSVWLLRLAKTYETTTLEK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYETARRHPDYSVWLLRLAKTYETTTLEK 384
QY 361 CAADAPHECVAKYVDFEFPKPLVEEPONLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAADAPHECVAKYVDFEFPKPLVEEPONLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 10
US-10-153-064-7
; Sequence 7, Application US/10153064
; Publication No. US200201428141
; GENERAL INFORMATION:
; APPLICANT: Bell et al.
; TITLE OF INVENTION: Chemokine Beta-1 Fusion Proteins
; FILE REFERENCE: PF556
; CURRENT APPLICATION NUMBER: US/10/153,064
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/293,212
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-153-064-7

Query Match 100.0%; Score 3103; DB 14; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFKDGLGEENFKALVLIAPAFQYLQCCPPEDHVLVNEVTEFAKTCVADSAE 60
Db 25 DAHSEVAHRFKDGLGEENFKALVLIAPAFQYLQCCPPEDHVLVNEVTEFAKTCVADSAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPLRVPEV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKOEPRNECFLOHKDDNPNLPLRVPEV 144
QY 121 DVMTAFPHDNEETFLKKLYIETARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 180
Db 145 DVMTAFPHDNEETFLKKLYIETARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 204
QY 181 KLDELDEGKASAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVS KLVDLTJK 240
Db 205 KLDELDEGKASAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVS KLVDLTJK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYETARRHPDYSVWLLRLAKTYETTTLEK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYETARRHPDYSVWLLRLAKTYETTTLEK 384
QY 361 CAADAPHECVAKYVDFEFPKPLVEEPONLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAADAPHECVAKYVDFEFPKPLVEEPONLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 609
```

```
Db 145 DVMTAFPHDNEETFLKKLYIETARRHPYFYAPELLFFAKRYKAATTECCQAADKAACLLP 204
QY 181 KLDELDEGKASAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVS KLVDLTJK 240
Db 205 KLDELDEGKASAKORLKASLOKFGERAFKAWAVARLSORFPKAEFAEVS KLVDLTJK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
Db 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYETARRHPDYSVWLLRLAKTYETTTLEK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYETARRHPDYSVWLLRLAKTYETTTLEK 384
QY 361 CAADAPHECVAKYVDFEFPKPLVEEPONLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 420
Db 385 CAADAPHECVAKYVDFEFPKPLVEEPONLIKONCELFPOLGEYKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 480
Db 445 PTLVEVSRLNGKVGSKCKHPKAEKMPCAEDYLSVVLNOLCVLHEKTPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKEROIKKQATLVELVKKPKAT 564
QY 541 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDDDFAAFVEKCKADDKETCFABEGKKLVAASQAALGL 609

RESULT 11
US-09-984-186-2
; Sequence 2, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fieer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
ADDRESS: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3043
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
```

REFERENCE/DOCKET NUMBER: S92006-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3839

TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-984-186-2

Query Match 100.0%; Score 3103; DB 10; Length 610;

Best Local Similarity 100.0%; Pred. No. 1.1e-270;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 60

Db 25 DAHSEVAHFKDGLGENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 84

QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPVLRPEV 120

Db 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPVLRPEV 144

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204

QY 181 KLDELDEGKASSAKORLKCSLQKGFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240

Db 205 KLDELDEGKASSAKORLKCSLQKGFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 264

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYENDEMPA 300

Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVILLRLAKTYETTLK 360

Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVILLRLAKTYETTLK 384

QY 361 CAADPHCEYAKYDFDEKPLVEEPONLIKQNCHELFQOLGEYKFQNALVRYTKKPVQST 420

Db 385 CAADPHCEYAKYDFDEKPLVEEPONLIKQNCHELFQOLGEYKFQNALVRYTKKPVQST 444

QY 421 PLYEVSRLNGKYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 445 PLYEVSRLNGKYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRRPCFSALEVDYVYKPENAETFTFHADICTLSEKROIKKQOTALVELVKHKPKAT 540

Db 505 LVNRRPCFSALEVDYVYKPENAETFTFHADICTLSEKROIKKQOTALVELVKHKPKAT 564

QY 541 KEQLKAVADDFAAFVEKCKKADDKETCFAPEGKKLVAAASQAALGL 585

Db 565 KEQLKAVADDFAAFVEKCKKADDKETCFAPEGKKLVAAASQAALGL 609

RESULT 12

US-10-237-667-2

Sequence 2, Application US/10237667

Publication No. US2003002308A1

GENERAL INFORMATION:

APPLICANT: Fleet

Fournier, Alain

Guittion, Jean-Dominique

Jung, Gerard

Reh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer Inc.

STREET: 500 Arcola Road, 3C43

CITY: Collegeville

STATE: PA

COUNTRY: USA

ZIP: 19426

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: System 7.1

SOFTWARE: Word 5.1 (patentin)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/237,667

FILING DATE: 10-Sep-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/797,689

FILING DATE: 31-JAN-1997

APPLICATION NUMBER: US 08/256,927

FILING DATE: 28-JUL-1994

APPLICATION NUMBER: FR 92/01064

FILING DATE: 31-JAN-1992

APPLICATION NUMBER: PCT/FR93/00085

FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Smith Ph.D., Julie K.

REGISTRATION NUMBER: P-38,619

REFERENCE/DOCKET NUMBER: S92006-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (610) 454-3839

TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 610 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-237-667-2

Query Match

Best Local Similarity 100.0%; Score 3103; DB 15; Length 610;

Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 60

Db 25 DAHSEVAHFKDGLGENFKALVIAFAQYLQCCPEDHVKLVNVEVTEFAKTCVADESAE 84

QY 61 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPVLRPEV 120

Db 85 NCKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPVLRPEV 144

QY 121 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 180

Db 145 DVMCTAFHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAFTCCQAADKAAACLLP 204

QY 181 KLDELDEGKASSAKORLKCSLQKGFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 240

Db 205 KLDELDEGKASSAKORLKCSLQKGFGERAFKAWAVARLSQRFPAEFAEVSKLVTDLTK 264

QY 241 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYENDEMPA 300

Db 265 VHTCCCHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEYENDEMPA 324

QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVILLRLAKTYETTLK 360

Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLMFLEYARRHPDYSVILLRLAKTYETTLK 384

QY 361 CAADPHCEYAKYDFDEKPLVEEPONLIKQNCHELFQOLGEYKFQNALVRYTKKPVQST 420

Db 385 CAADPHCEYAKYDFDEKPLVEEPONLIKQNCHELFQOLGEYKFQNALVRYTKKPVQST 444

QY 421 PLYEVSRLNGKYGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPKCAEDYLSVNLQCVLHKTVPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDEYVVKFNATFTFFHADICTLSEKEROIKKOTALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDEYVVKFNATFTFFHADICTLSEKEROIKKOTALVELVHKPKAT 564
QY 541 KEQLKAVMDFFAAFEVKCKADKCTCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDFFAAFEVKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 13
US-10-237-708-2
; Sequence 2, Application US/10237708
; Publication No. US20030036170A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guignon, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,708
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-39,619
; REFERENCE/DOCKET NUMBER: S192006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-708-2
Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFKDGENFKALVLIATFAQYLQCPDFHVKLVNEVTEFAKTCVADESAAE 60
|||||

Db 25 DAHSEVAHFKDGENFKALVLIATFAQYLQCPDFHVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCDLSLHTLFGDKLCVATLRETYGEMADCCAKOPEPERNECFLOHKDDNPNLPLRVPEV 120
Db 85 NCDLSLHTLFGDKLCVATLRETYGEMADCCAKOPEPERNECFLOHKDDNPNLPLRVPEV 144
QY 121 DVMCTAFHDNEETFLKKLYLIEIARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACILP 180
Db 145 DVMCTAFHDNEETFLKKLYLIEIARHPYFYAPPELLFFAKRYKAAFTCCQADKAAACILP 204
QY 181 KIDELRDEGKASSAKORLKCASLOKFGRAFKAWAVARLSORPPKRAEYFSKLYTDLTK 240
Db 205 KIDELRDEGKASSAKORLKCASLOKFGRAFKAWAVARLSORPPKRAEYFSKLYTDLTK 264
QY 241 VTECHCHDLLECADRADLAKYICENODSISKKLCECKPILLESKSHCIAEVENDEMPA 300
Db 265 VTECHCHDLLECADRADLAKYICENODSISKKLCECKPILLESKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARHPDYSVLLRLAKTYETTLK 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFLGMFLYIYARHPDYSVLLRLAKTYETTLK 384
QY 361 CAADPHECTAKYVDFEKLPLVEEPONLLKONCELFEOLEKFNALLVRYTKKVPQVST 420
Db 385 CAADPHECTAKYVDFEKLPLVEEPONLLKONCELFEOLEKFNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRNLGKVGSKCKHPKAPKMPKCAEDYLSVNLQCVLHKTVPVSDRVTKCTES 480
Db 445 PTLVEVSRNLGKVGSKCKHPKAPKMPKCAEDYLSVNLQCVLHKTVPVSDRVTKCTES 504
QY 481 LVNRRPCFSALEVDEYVVKFNATFTFFHADICTLSEKEROIKKOTALVELVHKPKAT 540
Db 505 LVNRRPCFSALEVDEYVVKFNATFTFFHADICTLSEKEROIKKOTALVELVHKPKAT 564
QY 541 KEQLKAVMDFFAAFEVKCKADKCTCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAVMDFFAAFEVKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 14
US-10-237-866-2
; Sequence 2, Application US/10237866
; Publication No. US20030036171A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guignon, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,866
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-39,619
; REFERENCE/DOCKET NUMBER: S192006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-866-2
Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DAHSEVAHFKDGENFKALVLIATFAQYLQCPDFHVKLVNEVTEFAKTCVADESAAE 60
|||||

FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-866-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESSE 60
Db 25 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESSE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 144
QY 121 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSKCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSKCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGFLYIYARRHPDYSVLLLLAKTYETILEKC 360
Db 325 DLPSLAADFVESKDVCKNYAEAKDVFGLGFLYIYARRHPDYSVLLLLAKTYETILEKC 384
QY 361 CAADPHCEYAKVDFDEKPLVEEPONLIKQNCLEFQGLGEYKFQNALIVRYTKKVPQVST 420
Db 385 CAADPHCEYAKVDFDEKPLVEEPONLIKQNCLEFQGLGEYKFQNALIVRYTKKVPQVST 444
QY 421 PTLVEVRNKGKVGSKCKHPKAMPKPCADYLSVNLNOLCVLHKEKTPVSDRVTKCCTES 480
Db 445 PTLVEVRNKGKVGSKCKHPKAMPKPCADYLSVNLNOLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVDYVPEKFNATFTFHADICTLSEKEROIKKQTFALVELVKKPKAT 540
Db 505 LVNRRPCFSALEVDYVPEKFNATFTFHADICTLSEKEROIKKQTFALVELVKKPKAT 564
QY 541 KEQLKAYMDFAAFVEKCKADDDKTCFAEEGKKLVAASQAALGL 585
Db 565 KEQLKAYMDFAAFVEKCKADDDKTCFAEEGKKLVAASQAALGL 609

RESULT 15
US-10-237-871-2
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain

Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3c43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,871
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-237-871-2

Query Match 100.0%; Score 3103; DB 15; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESSE 60
Db 25 DAHSEVAHFRKDLGEENFKALVLIAPAFYLOQCPEDHVKLVNVEVTEFAKTCVADESSE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 120
Db 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDNNPLRLVPRV 144
QY 121 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 180
Db 145 DVMCTAFPHDNEETFLKKYLYEIARRHPYFYAPPELLFFAKRYKAAPTECCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 240
Db 205 KLDELDEGKASSAKQRLKCSLOKFGERAFKAWARLSQRFPAKFAEVSKLIVDLTK 264
QY 241 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSKCIAEVENDEMPA 300
Db 265 VHTCCCHGDLLECCADRADLAKYICENQDSISSKLEKCEKPLLEKSKCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAEAKDVFGLGFLYIYARRHPDYSVLLLLAKTYETILEKC 360

```
|||||
Db 325 DLPSLAADFVESKDVCKNVAEKDVFLGMFLYBYARRHPDYSVLLLELAKYETITLEKC 384
QY 361 CAAADPHECYAKVDFEFLVEEPQNLKONCELFQOLGEYKFQNALLVRYTKKVPQVST 420
Db 385 CAAADPHECYAKVDFEFLVEEPQNLKONCELFQOLGEYKFQNALLVRYTKKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPKMPKPCARDYLSVVNLQLCVLHEKTPVSDRVTKCCTES 480
Db 445 PTLVEVSRLGKVGSKCKHPKMPKPCARDYLSVVNLQLCVLHEKTPVSDRVTKCCTES 504
QY 481 LVNRRPCFSALEVEDEYVPKEFNATFTTFHADICTLSEKEROIKKQOTALVELYKHKPKAT 540
Db 505 LVNRRPCFSALEVEDEYVPKEFNATFTTFHADICTLSEKEROIKKQOTALVELYKHKPKAT 564
QY 541 KEQLKAYMDDFAAFVEKCKCKADDDKETCFABEGKKLVAASQAALGL 585
Db 565 KEQLKAYMDDFAAFVEKCKCKADDDKETCFABEGKKLVAASQAALGL 609
```

Search completed: August 31, 2003, 16:39:14
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 43 Seconds
(without alignments)
1308.341 Million cell updates/sec

Title: US-09-833-041-18
Perfect score: 3103
Sequence: 1 DAHKSEVAHFRKDLGEENFK.....TCFAEKGKLVAAASQAALGL 585

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3103	100.0	609	1 ABHUS	serum albumin prec
2	2942	94.8	600	2 A47391	serum albumin prec
3	2620	84.4	608	2 S57632	serum albumin prec
4	2475.5	79.8	607	1 ABHOS	serum albumin prec
5	2446.5	78.8	607	1 ABHOS	serum albumin prec
6	2432.5	78.4	607	1 ABHOS	serum albumin prec
7	2426	78.2	608	1 ABHOS	serum albumin prec
8	2411.5	77.7	605	1 ABHOS	serum albumin prec
9	2387	76.9	609	2 JC5838	albumin - Mongolia
10	1861	60.0	453	2 A05139	serum albumin - mo
11	1557.5	50.2	615	1 ABHOS	serum albumin prec
12	1253.5	40.4	609	2 JC4258	alpha-fetoprotein
13	1249.5	40.3	609	1 FEPU	alpha-fetoprotein
14	1242.5	40.0	609	1 FEPU	alpha-fetoprotein
15	1205	38.8	607	1 ABX172	74K albumin prec
16	1181.5	38.1	265	2 I46986	albumin - dog (fra
17	1175.5	37.9	608	1 ABX168	68K serum albumin
18	1084	34.9	605	1 FPMS	alpha-fetoprotein
19	1067	34.4	611	1 FPPT	alpha-fetoprotein
20	1055	34.0	599	1 A54906	afamin precursor -
21	928.5	29.9	614	2 S59517	serum albumin prec
22	928	29.9	608	2 A53195	afamin precursor -
23	747.5	24.1	608	1 ABONS1	serum albumin 1 pr
24	742.5	23.9	608	1 ABONS2	serum albumin 2 pr
25	699	22.5	382	2 A37253	serum albumin - bu
26	440.5	14.2	1423	1 S27941	serum albumin - se
27	386	12.4	474	1 VYHSD	vitamin D-binding
28	385	12.4	476	1 VYRTD	vitamin D-binding
29	372	12.0	472	1 A35327	vitamin D-binding

30 184 5.9 1819 2 A71928 cag island protein
31 184 5.9 1927 2 G64585 cag pathogenicity
32 141.5 4.6 1560 2 T30282 calcium-binding pr
33 137.5 4.4 1348 2 AG2558 hypothetical prote
34 134.5 4.3 1004 2 JC2221 major surface glyco
35 134 4.3 1083 2 JC2300 cell surface glyco
36 134 4.3 1780 2 T17272 hypothetical prote
37 133.5 4.3 1390 2 S51364 sperm tail-specifi
38 132.5 4.3 1070 2 T06733 kinesin homolog F2
39 131 4.2 1076 2 JC2217 major surface glyco
40 130 4.2 1175 2 D35815 myosin heavy chain
41 129 4.2 1017 2 JC4035 cell-cycle-depende
42 128.5 4.1 1051 2 PC4091 glycoprotein A - p
43 128 4.1 1005 2 A64465 hypothetical prote
44 128 4.1 3225 2 I52300 gliadin - human
45 126.5 4.1 1927 2 A59236 embryonic muscle m

ALIGNMENTS

RESULT 1

ABHUS
serum albumin precursor [validated] - human
N:Alternate names: preproalbumin
N:Contains: kinetensin
C:Species: Homo sapiens (man)
C:Date: 29-Jul-1981 #sequence_revision 31-Jan-1997 #text_change 17-Mar-2000
C:Accession: A93743; A93936; I39427; I59286; I59313; G01747; S55314; A91420; S06422;
R:Lawn, R.M.; Adelman, J.; Bock, S.C.; Franke, C.M.; Houch, C.M.; Najarian, R.C.; Se
Nucleic Acids Res. 9, 6103-6114, 1981
A:Title: The sequence of human serum albumin cDNA and its expression in Escherichia
A:Reference number: A93743; MUID:82081882; PMID:6171778
A:Accession: A93743
A:Molecule type: mRNA
A:Residues: 1-419, 'K', 421-609 <LAW>
A:Cross-references: EMBL:Y00494; GB:J000078; GB:L00132; GB:L00133; NID:g28591; PIDN: C
R:Dugaiczky, A.; Law, S.W.; Dennison, O.E.
Proc. Natl. Acad. Sci. U.S.A. 79, 71-75, 1982
A:Title: Nucleotide sequence and the encoded amino acids of human serum albumin mRNA
A:Reference number: A93936; MUID:82105994; PMID:6273391
A:Accession: A93936
A:Molecule type: mRNA
A:Residues: 1-120, 'G', 122-609 <DUG>
A:Cross-references: EMBL:Y00494; NID:g28589; PIDN:CRAA23753.1; PID:g28590
R:Grano, Y.; Watanabe, K.; Sakai, M.; Tamaoki, T.
J. Biol. Chem. 261, 3244-3251, 1986
A:Title: The human albumin gene. Characterization of the 5' and 3' flanking regions
A:Reference number: I39427; MUID:86140099; PMID:2419329
A:Accession: I39427
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-26 <URA>
A:Cross-references: GB:M13075; NID:g178330; PIDN:AAA51688.1; PID:g553173
R:Watkins, S.; Madison, J.; Galliano, M.; Minchiotti, L.; Putnam, F.W.
Proc. Natl. Acad. Sci. U.S.A. 91, 2275-2279, 1994
A:Title: A nucleotide insertion and frameshift cause analbuminemia in an Italian fam
A:Reference number: I59286; MUID:94181575; PMID:8134387
A:Accession: I59286
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 282-290, 'KSRFDLO', <WAT>
A:Cross-references: GB:S69192; NID:g546032; PIDN:AAB30282.1; PID:g546033
A:Note: this frame-shift variant, designated albumin Roma, leads to analbuminemia
R:Madison, J.; Galliano, M.; Watkins, S.; Minchiotti, L.; Porta, F.; Rossi, A.; Putn
Proc. Natl. Acad. Sci. U.S.A. 91, 6476-6480, 1994
A:Title: Genetic variants of human serum albumin in Italy: point mutants and a carb
A:Reference number: I59313; MUID:94294404; PMID:8022807
A:Accession: I59313
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 589-590, 'ALPRVRKLLQVQLP', <MAD>
A:Cross-references: GB:S70799; NID:g547231; PIDN:AAB31177.1; PID:g547232

A;Note: this frame-shift variant is designated albumin Bazzano; four additional variants
 R;Menaya, J.; Parrilla, R.; Ayuso, M.S.
 Submitted to the EMBL Data Library, March 1995
 A;Reference number: G08292
 A;Accession: G01747
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-120, 'G', 122-455 <MEN>
 A;Cross-references: EMBL:U22961; NID:9763428; PID:AAA64922.1; PID:9763431
 R;Ledgerwood, E.C.; George, P.M.; Peach, R.J.; Brennan, S.O.
 Biochem. J. 308, 321-325, 1995
 A;Title: Endoproteolytic processing of recombinant proalbumin variants by the yeast *Kex2*
 A;Reference number: S55314; MUID:95275251; PMID:7755581
 A;Accession: S55314
 A;Molecule type: protein
 A;Residues: 19-27 <LED>
 R;Meioun, B.; Moravsek, L.; Kostka, V.
 FEBS Lett. 58, 134-137, 1975
 A;Title: Complete amino acid sequence of human serum albumin.
 A;Reference number: A91420; MUID:76187907; PMID:1225573
 A;Accession: A91420
 A;Molecule type: protein
 A;Residues: 25-117, 'Q', 120-154, 'Q', 156-193, 'E', 195-387, 'H', 389-390, 'Y', 392-393, 'A', 395-396, 'D', 398-400, 'G', 402-403, 'E', 405-406, 'D', 408-409, 'E', 411-412, 'E', 414-415, 'E', 417-418, 'E', 420-421, 'E', 423-424, 'E', 426-427, 'E', 429-430, 'E', 432-433, 'E', 435-436, 'E', 438-439, 'E', 441-442, 'E', 444-445, 'E', 447-448, 'E', 450-451, 'E', 453-454, 'E', 456-457, 'E', 459-460, 'E', 462-463, 'E', 465-466, 'E', 468-469, 'E', 471-472, 'E', 474-475, 'E', 477-478, 'E', 480-481, 'E', 483-484, 'E', 486-487, 'E', 489-490, 'E', 492-493, 'E', 495-496, 'E', 498-499, 'E', 501-502, 'E', 504-505, 'E', 507-508, 'E', 510-511, 'E', 513-514, 'E', 516-517, 'E', 519-520, 'E', 522-523, 'E', 525-526, 'E', 528-529, 'E', 531-532, 'E', 534-535, 'E', 537-538, 'E', 540-541, 'E', 543-544, 'E', 546-547, 'E', 549-550, 'E', 552-553, 'E', 555-556, 'E', 558-559, 'E', 561-562, 'E', 564-565, 'E', 567-568, 'E', 570-571, 'E', 573-574, 'E', 576-577, 'E', 579-580, 'E', 582-583, 'E', 585-586, 'E', 588-589, 'E', 591-592, 'E', 594-595, 'E', 597-598, 'E', 600-601, 'E', 603-604, 'E', 606-607, 'E', 609-610, 'E', 612-613, 'E', 615-616, 'E', 618-619, 'E', 621-622, 'E', 623-624, 'E', 626-627, 'E', 629-630, 'E', 632-633, 'E', 635-636, 'E', 638-639, 'E', 641-642, 'E', 643-644, 'E', 646-647, 'E', 649-650, 'E', 652-653, 'E', 655-656, 'E', 658-659, 'E', 661-662, 'E', 664-665, 'E', 667-668, 'E', 670-671, 'E', 673-674, 'E', 676-677, 'E', 679-680, 'E', 682-683, 'E', 685-686, 'E', 688-689, 'E', 691-692, 'E', 694-695, 'E', 697-698, 'E', 700-701, 'E', 703-704, 'E', 706-707, 'E', 709-710, 'E', 712-713, 'E', 715-716, 'E', 718-719, 'E', 721-722, 'E', 724-725, 'E', 727-728, 'E', 730-731, 'E', 733-734, 'E', 736-737, 'E', 739-740, 'E', 742-743, 'E', 745-746, 'E', 748-749, 'E', 751-752, 'E', 754-755, 'E', 757-758, 'E', 760-761, 'E', 763-764, 'E', 766-767, 'E', 769-770, 'E', 772-773, 'E', 775-776, 'E', 778-779, 'E', 781-782, 'E', 784-785, 'E', 787-788, 'E', 790-791, 'E', 793-794, 'E', 796-797, 'E', 799-800, 'E', 802-803, 'E', 805-806, 'E', 808-809, 'E', 811-812, 'E', 814-815, 'E', 817-818, 'E', 820-821, 'E', 823-824, 'E', 826-827, 'E', 829-830, 'E', 832-833, 'E', 835-836, 'E', 838-839, 'E', 841-842, 'E', 843-844, 'E', 846-847, 'E', 849-850, 'E', 852-853, 'E', 855-856, 'E', 858-859, 'E', 861-862, 'E', 864-865, 'E', 867-868, 'E', 870-871, 'E', 873-874, 'E', 876-877, 'E', 879-880, 'E', 882-883, 'E', 885-886, 'E', 888-889, 'E', 891-892, 'E', 894-895, 'E', 897-898, 'E', 900-901, 'E', 903-904, 'E', 906-907, 'E', 909-910, 'E', 912-913, 'E', 915-916, 'E', 918-919, 'E', 921-922, 'E', 924-925, 'E', 927-928, 'E', 930-931, 'E', 933-934, 'E', 936-937, 'E', 939-940, 'E', 942-943, 'E', 945-946, 'E', 948-949, 'E', 951-952, 'E', 954-955, 'E', 957-958, 'E', 960-961, 'E', 963-964, 'E', 966-967, 'E', 969-970, 'E', 972-973, 'E', 975-976, 'E', 978-979, 'E', 981-982, 'E', 984-985, 'E', 987-988, 'E', 990-991, 'E', 993-994, 'E', 996-997, 'E', 999-1000, 'E', 1002-1003, 'E', 1005-1006, 'E', 1008-1009, 'E', 1011-1012, 'E', 1014-1015, 'E', 1017-1018, 'E', 1020-1021, 'E', 1023-1024, 'E', 1026-1027, 'E', 1029-1030, 'E', 1032-1033, 'E', 1035-1036, 'E', 1038-1039, 'E', 1041-1042, 'E', 1043-1044, 'E', 1046-1047, 'E', 1049-1050, 'E', 1052-1053, 'E', 1055-1056, 'E', 1058-1059, 'E', 1061-1062, 'E', 1064-1065, 'E', 1067-1068, 'E', 1070-1071, 'E', 1073-1074, 'E', 1076-1077, 'E', 1079-1080, 'E', 1082-1083, 'E', 1085-1086, 'E', 1088-1089, 'E', 1091-1092, 'E', 1094-1095, 'E', 1097-1098, 'E', 1100-1101, 'E', 1103-1104, 'E', 1106-1107, 'E', 1109-1110, 'E', 1112-1113, 'E', 1115-1116, 'E', 1118-1119, 'E', 1121-1122, 'E', 1124-1125, 'E', 1127-1128, 'E', 1130-1131, 'E', 1133-1134, 'E', 1136-1137, 'E', 1139-1140, 'E', 1142-1143, 'E', 1145-1146, 'E', 1148-1149, 'E', 1151-1152, 'E', 1154-1155, 'E', 1157-1158, 'E', 1160-1161, 'E', 1163-1164, 'E', 1166-1167, 'E', 1169-1170, 'E', 1172-1173, 'E', 1175-1176, 'E', 1178-1179, 'E', 1181-1182, 'E', 1184-1185, 'E', 1187-1188, 'E', 1190-1191, 'E', 1193-1194, 'E', 1196-1197, 'E', 1199-1200, 'E', 1202-1203, 'E', 1205-1206, 'E', 1208-1209, 'E', 1211-1212, 'E', 1214-1215, 'E', 1217-1218, 'E', 1220-1221, 'E', 1223-1224, 'E', 1226-1227, 'E', 1229-1230, 'E', 1232-1233, 'E', 1235-1236, 'E', 1238-1239, 'E', 1241-1242, 'E', 1243-1244, 'E', 1246-1247, 'E', 1249-1250, 'E', 1252-1253, 'E', 1255-1256, 'E', 1258-1259, 'E', 1261-1262, 'E', 1264-1265, 'E', 1267-1268, 'E', 1270-1271, 'E', 1273-1274, 'E', 1276-1277, 'E', 1279-1280, 'E', 1282-1283, 'E', 1285-1286, 'E', 1288-1289, 'E', 1291-1292, 'E', 1294-1295, 'E', 1297-1298, 'E', 1300-1301, 'E', 1303-1304, 'E', 1306-1307, 'E', 1309-1310, 'E', 1312-1313, 'E', 1315-1316, 'E', 1318-1319, 'E', 1321-1322, 'E', 1324-1325, 'E', 1327-1328, 'E', 1330-1331, 'E', 1333-1334, 'E', 1336-1337, 'E', 1339-1340, 'E', 1342-1343, 'E', 1345-1346, 'E', 1348-1349, 'E', 1351-1352, 'E', 1354-1355, 'E', 1357-1358, 'E', 1360-1361, 'E', 1363-1364, 'E', 1366-1367, 'E', 1369-1370, 'E', 1372-1373, 'E', 1375-1376, 'E', 1378-1379, 'E', 1381-1382, 'E', 1384-1385, 'E', 1387-1388, 'E', 1390-1391, 'E', 1393-1394, 'E', 1396-1397, 'E', 1399-1400, 'E', 1402-1403, 'E', 1405-1406, 'E', 1408-1409, 'E', 1411-1412, 'E', 1414-1415, 'E', 1417-1418, 'E', 1420-1421, 'E', 1423-1424, 'E', 1426-1427, 'E', 1429-1430, 'E', 1432-1433, 'E', 1435-1436, 'E', 1438-1439, 'E', 1441-1442, 'E', 1443-1444, 'E', 1446-1447, 'E', 1449-1450, 'E', 1452-1453, 'E', 1455-1456, 'E', 1458-1459, 'E', 1461-1462, 'E', 1464-1465, 'E', 1467-1468, 'E', 1470-1471, 'E', 1473-1474, 'E', 1476-1477, 'E', 1479-1480, 'E', 1482-1483, 'E', 1485-1486, 'E', 1488-1489, 'E', 1491-1492, 'E', 1494-1495, 'E', 1497-1498, 'E', 1500-1501, 'E', 1503-1504, 'E', 1506-1507, 'E', 1509-1510, 'E', 1512-1513, 'E', 1515-1516, 'E', 1518-1519, 'E', 1521-1522, 'E', 1524-1525, 'E', 1527-1528, 'E', 1530-1531, 'E', 1533-1534, 'E', 1536-1537, 'E', 1539-1540, 'E', 1542-1543, 'E', 1545-1546, 'E', 1548-1549, 'E', 1551-1552, 'E', 1554-1555, 'E', 1557-1558, 'E', 1560-1561, 'E', 1563-1564, 'E', 1566-1567, 'E', 1569-1570, 'E', 1572-1573, 'E', 1575-1576, 'E', 1578-1579, 'E', 1581-1582, 'E', 1584-1585, 'E', 1587-1588, 'E', 1590-1591, 'E', 1593-1594, 'E', 1596-1597, 'E', 1599-1600, 'E', 1602-1603, 'E', 1605-1606, 'E', 1608-1609, 'E', 1611-1612, 'E', 1614-1615, 'E', 1617-1618, 'E', 1620-1621, 'E', 1623-1624, 'E', 1626-1627, 'E', 1629-1630, 'E', 1632-1633, 'E', 1635-1636, 'E', 1638-1639, 'E', 1641-1642, 'E', 1643-1644, 'E', 1646-1647, 'E', 1649-1650, 'E', 1652-1653, 'E', 1655-1656, 'E', 1658-1659, 'E', 1661-1662, 'E', 1664-1665, 'E', 1667-1668, 'E', 1670-1671, 'E', 1673-1674, 'E', 1676-1677, 'E', 1679-1680, 'E', 1682-1683, 'E', 1685-1686, 'E', 1688-1689, 'E', 1691-1692, 'E', 1694-1695, 'E', 1697-1698, 'E', 1700-1701, 'E', 1703-1704, 'E', 1706-1707, 'E', 1709-1710, 'E', 1712-1713, 'E', 1715-1716, 'E', 1718-1719, 'E', 1721-1722, 'E', 1724-1725, 'E', 1727-1728, 'E', 1730-1731, 'E', 1733-1734, 'E', 1736-1737, 'E', 1739-1740, 'E', 1742-1743, 'E', 1745-1746, 'E', 1748-1749, 'E', 1751-1752, 'E', 1754-1755, 'E', 1757-1758, 'E', 1760-1761, 'E', 1763-1764, 'E', 1766-1767, 'E', 1769-1770, 'E', 1772-1773, 'E', 1775-1776, 'E', 1778-1779, 'E', 1781-1782, 'E', 1784-1785, 'E', 1787-1788, 'E', 1790-1791, 'E', 1793-1794, 'E', 1796-1797, 'E', 1799-1800, 'E', 1802-1803, 'E', 1805-1806, 'E', 1808-1809, 'E', 1811-1812, 'E', 1814-1815, 'E', 1817-1818, 'E', 1820-1821, 'E', 1823-1824, 'E', 1826-1827, 'E', 1829-1830, 'E', 1832-1833, 'E', 1835-1836, 'E', 1838-1839, 'E', 1841-1842, 'E', 1843-1844, 'E', 1846-1847, 'E', 1849-1850, 'E', 1852-1853, 'E', 1855-1856, 'E', 1858-1859, 'E', 1861-1862, 'E', 1864-1865, 'E', 1867-1868, 'E', 1870-1871, 'E', 1873-1874, 'E', 1876-1877, 'E', 1879-1880, 'E', 1882-1883, 'E', 1885-1886, 'E', 1888-1889, 'E', 1891-1892, 'E', 1894-1895, 'E', 1897-1898, 'E', 1900-1901, 'E', 1903-1904, 'E', 1906-1907, 'E', 1909-1910, 'E', 1912-1913, 'E', 1915-1916, 'E', 1918-1919, 'E', 1921-1922, 'E', 1924-1925, 'E', 1927-1928, 'E', 1930-1931, 'E', 1933-1934, 'E', 1936-1937, 'E', 1939-1940, 'E', 1942-1943, 'E', 1945-1946, 'E', 1948-1949, 'E', 1951-1952, 'E', 1954-1955, 'E', 1957-1958, 'E', 1960-1961, 'E', 1963-1964, 'E', 1966-1967, 'E', 1969-1970, 'E', 1972-1973, 'E', 1975-1976, 'E', 1978-1979, 'E', 1981-1982, 'E', 1984-1985, 'E', 1987-1988, 'E', 1990-1991, 'E', 1993-1994, 'E', 1996-1997, 'E', 1999-2000, 'E', 2002-2003, 'E', 2005-2006, 'E', 2008-2009, 'E', 2011-2012, 'E', 2014-2015, 'E', 2017-2018, 'E', 2020-2021, 'E', 2023-2024, 'E', 2026-2027, 'E', 2029-2030, 'E', 2032-2033, 'E', 2035-2036, 'E', 2038-2039, 'E', 2041-2042, 'E', 2043-2044, 'E', 2046-2047, 'E', 2049-2050, 'E', 2052-2053, 'E', 2055-2056, 'E', 2058-2059, 'E', 2061-2062, 'E', 2064-2065, 'E', 2067-2068, 'E', 2070-2071, 'E', 2073-2074, 'E', 2076-2077, 'E', 2079-2080, 'E', 2082-2083, 'E', 2085-2086, 'E', 2088-2089, 'E', 2091-2092, 'E', 2094-2095, 'E', 2097-2098, 'E', 2100-2101, 'E', 2103-2104, 'E', 2106-2107, 'E', 2109-2110, 'E', 2112-2113, 'E', 2115-2116, 'E', 2118-2119, 'E', 2121-2122, 'E', 2124-2125, 'E', 2127-2128, 'E', 2130-2131, 'E', 2133-2134, 'E', 2136-2137, 'E', 2139-2140, 'E', 2142-2143, 'E', 2145-2146, 'E', 2148-2149, 'E', 2151-2152, 'E', 2154-2155, 'E', 2157-2158, 'E', 2160-2161, 'E', 2163-2164, 'E', 2166-2167, 'E', 2169-2170, 'E', 2172-2173, 'E', 2175-2176, 'E', 2178-2179, 'E', 2181-2182, 'E', 2184-2185, 'E', 2187-2188, 'E', 2190-2191, 'E', 2193-2194, 'E', 2196-2197, 'E', 2199-2200, 'E', 2202-2203, 'E', 2205-2206, 'E', 2208-2209, 'E', 2211-2212, 'E', 2214-2215, 'E', 2217-2218, 'E', 2220-2221, 'E', 2223-2224, 'E', 2226-2227, 'E', 2229-2230, 'E', 2232-2233, 'E', 2235-2236, 'E', 2238-2239, 'E', 2241-2242, 'E', 2243-2244, 'E', 2246-2247, 'E', 2249-2250, 'E', 2252-2253, 'E', 2255-2256, 'E', 2258-2259, 'E', 2261-2262, 'E', 2264-2265, 'E', 2267-2268, 'E', 2270-2271, 'E', 2273-2274, 'E', 2276-2277, 'E', 2279-2280, 'E', 2282-2283, 'E', 2285-2286, 'E', 2288-2289, 'E', 2291-2292, 'E', 2294-2295, 'E', 2297-2298, 'E', 2300-2301, 'E', 2303-2304, 'E', 2306-2307, 'E', 2309-2310, 'E', 2312-2313, 'E', 2315-2316, 'E', 2318-2319, 'E', 2321-2322, 'E', 2324-2325, 'E', 2327-2328, 'E', 2330-2331, 'E', 2333-2334, 'E', 2336-2337, 'E', 2339-2340, 'E', 2342-2343, 'E', 2345-2346, 'E', 2348-2349, 'E', 2351-2352, 'E', 2354-2355, 'E', 2357-2358, 'E', 2360-2361, 'E', 2363-2364, 'E', 2366-2367, 'E', 2369-2370, 'E', 2372-2373, 'E', 2375-2376, 'E', 2378-2379, 'E', 2381-2382, 'E', 2384-2385, 'E', 2387-2388, 'E', 2390-2391, 'E', 2393-2394, 'E', 2396-2397, 'E', 2399-2400, 'E', 2402-2403, 'E', 2405-2406, 'E', 2408-2409, 'E', 2411-2412, 'E', 2414-2415, 'E', 2417-2418, 'E', 2420-2421, 'E', 2423-2424, 'E', 2426-2427, 'E', 2429-2430, 'E', 2432-2433, 'E', 2435-2436, 'E', 2438-2439, 'E', 2441-2442, 'E', 2443-2444, 'E', 2446-2447, 'E', 2449-2450, 'E', 2452-2453, 'E', 2455-2456, 'E', 2458-2459, 'E', 2461-2462, 'E', 2464-2465, 'E', 2467-2468, 'E', 2470-2471, 'E', 2473-2474, 'E', 2476-2477, 'E', 2479-2480, 'E', 2482-2483, 'E', 2485-2486, 'E', 2488-2489, 'E', 2491-2492, 'E', 2494-2495, 'E', 2497-2498, 'E', 2500-2501, 'E', 2503-2504, 'E', 2506-2507, 'E', 2509-2510, 'E', 2512-2513, 'E', 2515-2516, 'E', 2518-2519, 'E', 2521-2522, 'E', 2524-2525, 'E', 2527-2528, 'E', 2530-2531, 'E', 2533-2534, 'E', 2536-2537, 'E', 2539-2540, 'E', 2542-2543, 'E', 2545-2546, 'E', 2548-2549, 'E', 2551-2552, 'E', 2554-2555, 'E', 2557-2558, 'E', 2560-2561, 'E', 2563-2564, 'E', 2566-2567, 'E', 2569-2570, 'E', 2572-2573, 'E', 2575-2576, 'E', 2578-2579, 'E', 2581-2582, 'E', 2584-2585, 'E', 2587-2588, 'E', 2590-2591, 'E', 2593-2594, 'E', 2596-2597, 'E', 2599-2600, 'E', 2602-2603, 'E', 2605-2606, 'E', 2608-2609, 'E', 2611-2612, 'E', 2614-2615, 'E', 2617-2618, 'E', 2620-2621, 'E', 2623-2624, 'E', 2626-2627, 'E', 2629-2630, 'E', 2632-2633, 'E', 2635-2636, 'E', 2638-2639, 'E', 2641-2642, 'E', 2643-2644, 'E', 2646-2647, 'E', 2649-2650, 'E', 2652-2653, 'E', 2655-2656, 'E', 2658-2659, 'E', 2661-2662, 'E', 2664-2665, 'E', 2667-2668, 'E', 2670-2671, 'E', 2673-2674, 'E', 2676-2677, 'E', 2679-2680, 'E', 2682-2683, 'E', 2685-2686, 'E', 2688-2689, 'E', 2691-2692, 'E', 2694-2695, 'E', 2697-2698, 'E', 2700-2701, 'E', 2703-2704, 'E', 2706-2707, 'E', 2709-2710, 'E', 2712-2713, 'E', 2715-2716, 'E', 2718-2719, 'E', 2721-2722, 'E', 2724-2725, 'E', 2727-2728, 'E', 2730-2731, 'E', 2733-2734, 'E', 2736-2737, 'E', 2739-2740, 'E', 2742-2743, 'E', 2745-2746, 'E', 2748-2749, 'E', 2751-2752, 'E', 2754-2755, 'E', 2757-2758, 'E', 2760-2761, 'E', 2763-2764, 'E', 2766-2767, 'E', 2769-2770, 'E', 2772-2773, 'E', 2775-2776, 'E', 2778-2779, 'E', 2781-2782, 'E', 2784-2785, 'E', 2787-2788, 'E', 2790-2791, 'E', 2793-2794, 'E', 2796-2797, 'E', 2799-2800, 'E', 2802-2803, 'E', 2805-2806, 'E', 2808-2809, 'E', 2811-2812, 'E', 2814-2815, 'E', 2817-2818, 'E', 2820-2821, 'E', 2823-2824, 'E', 2826-2827, 'E', 2829-2830, 'E', 2832-2833, 'E', 2835-2836, 'E', 2838-2839, 'E', 2841-2842, 'E', 2843-2844, 'E', 2846-2847, 'E', 2849-2850, 'E', 2852-2853, 'E', 2855-2856, 'E', 2858-2859, 'E', 2861-2862, 'E', 2864-2865, 'E', 2867-2868, 'E', 2870-2871, 'E', 2873-2874, 'E', 2876-2877, 'E', 2879-2880, 'E', 2882-2883, 'E', 2885-2886, 'E', 2888-2889, 'E', 2891-2892, 'E', 2894-2895, 'E', 2897-2898, 'E', 2900-2901, 'E', 2903-2904, 'E', 2906-2907, 'E', 2909-2910, 'E', 2912-2913, 'E', 2915-2916, 'E', 2918-2919, 'E', 2921-2922, 'E', 2924-2925, 'E', 2927-2928, 'E', 2930-2931, 'E', 2933-2934, 'E', 2936-2937, 'E', 2939-2940, 'E', 2942-2943, 'E', 2945-2946, 'E', 2948-2949, 'E', 2951-2952, 'E', 2954-2955, 'E', 2957-2958, 'E', 2960-2961, 'E', 29

F;166-174/Product: kinetensin #status experimental <KIP>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA2>
F;27/Binding site: copper (His) #status predicted
F;77-86,99-115,114-125,148-193,192-201,224-270,269-303,302-313,340-385,384-393,413-592/Binding site: pyridoxal phosphate (lys) (covalent) #status experimental

Query Match 100.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 9,2e-198;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGEENFKALVLIAPAYLOQCFFEDHVKLVNVEFEAKTCVADESAE 60
DB 25 DAHSEVAHFKDGLGEENFKALVLIAPAYLOQCFFEDHVKLVNVEFEAKTCVADESAE 84
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 85 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 144
QY 121 DVMCTAFHNEETFLAKYIETARHHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180
DB 145 DVMCTAFHNEETFLAKYIETARHHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240
DB 205 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 264
QY 241 VHTCCGDLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 265 VHTCCGDLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYETARRHPDYSVLLLRKATYETITLKC 360
DB 325 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYETARRHPDYSVLLLRKATYETITLKC 384
QY 361 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQELGEYKFONALLVRYTKVPQVST 420
DB 385 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQELGEYKFONALLVRYTKVPQVST 444
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 445 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 504
QY 481 LVNRPCFSALEVDYVTPKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 540
DB 505 LVNRPCFSALEVDYVTPKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 564
QY 541 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKKLVAAASQAALGL 585
DB 565 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKKLVAAASQAALGL 609

RESULT 2
A47391
serum albumin precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A47391
B:Watkins, S.; Sakamoto, Y.; Madison, J.; Davis, E.; Smith, D.G.; Dmulet, J.; Putnam, F.
Proc. Natl. Acad. Sci. U.S.A. 90, 2409-2413, 1993
A:Title: cDNA and protein sequence of polymorphic macaque albumins that differ in biliru-
bin binding site
A:Reference number: A47391; MUID:93211971; PMID:8460152
A:Contents: B/B homozygote
A:Accession: A47391
A>Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-600 <WAT>
A:Cross-references: GB:M50463; NID:g342294; PIDN:AAA36906.1; PID:g342295
A:Experimental source: liver
A>Note: sequence extracted from NCBI backbone (NCBIN:128280, NCBIp:128281).
C:Superfamily: serum albumin; serum albumin repeat homology
F;21-194/Domain: serum albumin repeat homology <SA1>
F;213-386/Domain: serum albumin repeat homology <SA2>

F;405-584/Domain: serum albumin repeat homology <SA3>
Query Match 94.8%; Score 2942; DB 2; Length 600;
Best Local Similarity 93.5%; Pred. No. 4e-187;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHFKDGLGEENFKALVLIAPAYLOQCFFEDHVKLVNVEFEAKTCVADESAE 60
DB 17 DTHKSEVAHFKDGLGEENFKALVLIAPAYLOQCFFEDHVKLVNVEFEAKTCVADESAE 76
QY 61 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 120
DB 77 NDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLVRPEV 136
QY 121 DVMCTAFHNEETFLAKYIETARHHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 180
DB 137 DVMCTAFHNEETFLAKYIETARHHPYFAPPELLFFAKRYKAATFCCQAADKAACLLP 196
QY 181 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 240
DB 197 KLDELDEGKASSAKORLKASLOKGERAFKAWAVARLSORFPKAEAEVSKLVTDLTK 256
QY 241 VHTCCGDLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 300
DB 257 VHTCCGDLLECADRADLAKYICENODSISSKLKECCPELLEKSHCIAEVENDEMPA 316
QY 301 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYETARRHPDYSVLLLRKATYETITLKC 360
DB 317 DLPSLAADFVESKDVCKNVAEAKDVLGMFLYETARRHPDYSVLLLRKATYETITLKC 376
QY 361 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQELGEYKFONALLVRYTKVPQVST 420
DB 377 CAADPHECYAKVDFEFPLVEEPONLIKQNCLEFQELGEYKFONALLVRYTKVPQVST 436
QY 421 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 480
DB 437 PTLVEVSRLGKVGSKCKHPEAKRMPCAEDYLSVVLNQLCVLHKEKTPVSDRVTKCCTES 496
QY 481 LVNRPCFSALEVDYVTPKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 540
DB 497 LVNRPCFSALEVDYVTPKFNATFTFHADICTLSEKROIKKOTALVELVKKPKAT 556
QY 541 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKKLVAAASQAAL 583
DB 557 KEQLKAVMDPFAFVEKCKCKADDKTCFABEGKKLVAAASQAAL 599

RESULT 3
S57632
serum albumin precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Aug-1999
C:Accession: J04660; S57632
R:Hilger, C.; Grigioni, F.; Hentges, F.
Gene 169, 295-296, 1996
A:Title: Sequence of the gene encoding cat (Felis domesticus) serum albumin.
A:Reference number: J04660; MUID:96194824; PMID:8647469
A:Accession: J04660
A:Molecule type: mRNA
A:Residues: 1-608 <H12>
A:Cross-references: EMBL:X84842; NID:g886484; PIDN:CAA59279.1; PID:g886485
A:Experimental source: liver
C:Comment: This protein is the major protein component in plasma. It functions as a
ein has 35 conserved cysteine residues.
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: liver; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRP>
F;25-608/Product: serum albumin #status predicted <MAT>
F;29-202/Domain: serum albumin repeat homology <SA1>
F;221-394/Domain: serum albumin repeat homology <SA2>
F;413-592/Domain: serum albumin repeat homology <SA3>

Query Match 79.8%; Score 2475.5; DB 1; Length 607;
Best Local Similarity 76.3%; Pred. No. 3e-156;
Matches 445; Conservative 69; Mismatches 68; Indels 1; Gaps 1;

QY 1 DAHSEVAREKDIGEENFKALVLIAFAOYLQOCPEDHVKLVNVEYFEAKTCVADESAE 60
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 25 DTHKSETAHRNDLGEKHFGKLVLVAFSQYLQOCPEDHVKLVNVEYFEAKCAADESAE 84
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 61 NCDKSHTLTFLGDKLCTVATLRRETYGENADCCAKOEPRNECFLOAHKDNPMLRLVPEVP 120
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 85 NCDKSHTLTFLGDKLCTVATLRATYGELADCCQEPRNECFIATHKDDHPNLKPL - KPEP 143
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 121 DVMTAFPHDNEETFLKYLYEIARRHYFYAPPELLFAKKRYKAAPTCCQAADAACLAIP 180
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 144 DAOCAAFQEDPKFTLGKLYIVARRHYFYTGPELLFHAETHYKADFTTCCPADDKLACLIP 203
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 181 KLDELDEGKASSAKORLKASLOKFGERAFKAWAVARLSOFEPKABFAEVSKLIVDLTK 240
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 204 KLDAIKERILLSSAKERLUKSCSFONGERAKAWSVARLSQKFPADFVAEVSIVTDLT 263
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 241 VHTCECHGDLLCACDDRADLAKYICENQDSISSLKECCERPLLEKSHCIAFVENDEMPA 300
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 264 VHKECCHGDLLCACDDRADLAKYICEHQDSISGLKACODKPELLKSHCIAEYKEDDLPS 323
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 301 DLPSLAADFVESKDYCKNYBAKVDFIGMPLYEARHPDYSVLLRLAKTTTFTEPLEKC 360
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 324 DLPALAADFADKEICKYIKDAKOVFTLGTLYEYSRRHPDYSVSLILRIARTYEATLEKC 383
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 361 CAAADPHCEYCAKVFDEFKPLVEEPONLIKONCELFBOLGEYKFONALLVRYTKKVPQST 420
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 384 CAEADPPACYRTVFDQFLPVEEPKSVLKNKCOLFEVGEGVDFOQLIVRYTKKAQVST 443
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 421 PTLVEYSRNILGKVGSKCKHPEAKRMPACBAYLSVVINOLCVLHEKTTPVSDRVTKCTES 480
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 444 PTLVEIGRTLKGVRCKCLKPESRLPCSENHLALNALRLCVLHEKTPVSEKITKCTCDS 503
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 481 LVNRPCFSALEVDETYPYKFNENAETFFHADICTLSEKERQIKQKQALVELYKHKPKAT 540
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 504 LAERPFCFSALEDGGYVPKFEKATETFFHADICTLPEDEKIQKQSALAEVLKHKPKAT 563
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
QY 541 KEQLKAVMDFAAFVEKCKCKADDKETCFEAEGKKLVAASAQAL 583
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
DB 564 KEQLKTVLGNFSFAVAKCGREDKEACFAEGKPLVASSQAL 606
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :

RESULT 5
ABOS

N;Alternate names: 67K protein; preproalbumin
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 30-Sep-1993 #text_change 18-Aug-2000
C;Accession: A38885; A36401; A91259; B60808; S10780; D45800; A26693; A90309; A91458;
R;Holowachuk, E.W.; Stoltenberg, J.K.; Reed, R.G.; Peters Jr., T.
submitted to the EMBL Data Library, August 1991
A;Description: Bovine serum albumin: cDNA sequence and expression.
A;Reference number: A38885
A;Accession: A38885
A:Molecule type: mRNA
A;Residues: 1-607 <HOL>
A;Cross-references: EMBL:M73215
R;Hiroyama, K.; Akashi, S.; Furiya, M.; Fukuhara, K.
Biochem. Biophys. Res. Commun. 173, 639-646, 1990
A;Title: Rapid confirmation and revision of the primary structure of bovine serum all.
Eur. J. Biochem. 98, 477-485, 1979
A;Reference number: A36401; MWID:91083649; PMID:2260975
A;Accession: A36401
A:Molecule type: protein
A;Residues: 25-41; H'43-189, 'E', 191-213, 'T', 215-323, 'D', 325-393, 'TS', 396-607 <HIR>
R;Macgillivray, R.T.A.; Chung, D.W.; Davie, E.W.
A;Title: Biosynthesis of bovine plasma proteins in a cell-free system.
A;Reference number: A91259; MWID:80024278; PMID:488109
A;Accession: A91258
A:Molecule type: protein

Best Local Similarity 75.6%; Pred. No. 2.5e-154;
Matches 441; Conservative 71; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHSEVHRFDLGEENFKALVLIAFAOYLQCPDEHDVKLVNEVEFAKTCVADESAG 60
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||::||
25 DTHSEIATHRFDLGEBOFGKVLIASFQYLQCPDDEHVLYNELFEFAKTCVADESHA 84
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 61 NCDKSLHTFLGDKLCTVATLRETYGMADCAKQEPERNECFLOHKDDNPNLPRLVRPEW 120
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
85 GCEKSLHTFLGDLCKVASLRETVDMDACCEKQEPERNECFLSHKDDSDPLKL-KPDP 143
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 121 DVMCTAPHNDNEEFLKKYLYTARSHHYFVAPELLFFAKRYKAFTGCCAADAACAALLP 180
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
144 NYLCDFEKADKKFWGKYLYETARRHYTFYAPELLYANKYNGVQCCEARDKGACLLP 203
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 181 KIDELRDEGRASSAKORLKRCASLOKGERAFKAWAVARLSQFPKAEFVSCLVTDLTK 240
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
204 KIETMREKVLASSARGQLRCASIQKGERALKAWSVARLSQPKPAEFVEVTKLVTDLFK 263
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 241 VHTCCGGDLLLEGADRADLAKEICENQDSISKLECECEKPLEKSHCIAEVENDEMA 300
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
264 VRKCCGGDLLLEGADRADLAKEICDNQDTISSKLECCDCLEKSHCIAEVEKDPAIE 323
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 301 DLPSLAADFVESGVCKNYAKOVFLGMLTYVARRHPDYSVLLLRIAKTVETTIERC 360
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
324 NLPLPTADFADKDVCKNYQEAQDAFGSTFYESSRHPEYAVSLLRAKEYENTLEEC 383
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 361 CAADAPHECYAKVFDEFKPLVEEPONLIKONCELFPOLGEYKFONALLVRYTKVPQVST 420
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
384 CAKDDPHACYSTVFDLKLHLVDPONLIKONCDQFKEGLGEYGFQNALIVTRYTRKPQVST 443
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 421 PLTYEVSRNLGKVGSKCKHPKAPKPCADEYLSVNLNOLCVLHKEKTPSDRVTKCOTES 480
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
444 PLTYEVSRNLGKVGSKCKHPKAPKPCADEYLSVNLNOLCVLHKEKTPSDRVTKCOTES 503
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
QY 481 LVNRPPCFSALEYDEVTEYVPEKFNATFTFHADICTLSEKERQIKQTALVELYKHKPKAT 540
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
504 LVNRPPCFSALETPEYVYPKAFDELFTFHADICTLPDTEKQIKQTALVELYKHKPKAT 563
QY 541 XEOLKANVMDFFAFVFKCKCKADDKETCAPRGKKLVAAQSQAAL 583
DB :|::||::||::||::||::||::||::||::||::||::||::||::||::||
564 ESQLKTVMNFVAFVDRCKAADCKEACFVBPGLKVVYSTQIAL 606

RESULT 6
ABSES
serum albumin precursor - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 31-Dec-1993#sequence-revision 31-Dec-1993 #text-change 22-Jun-1999
C:Accession: S06936
R:Brown, W.M.; Dziegielewska, K.M.; Foreman, R.C.; Saunders, N.R.
Nucleic Acids Res. 17, 10495, 1989
A:Title: Nucleotide and deduced amino acid sequence of sheep serum albumin.
A:Reference number: S06936; MUID:90098888; PMID:2602160
A:Accession: S06936
A:Molecule type: cDNA
A:Residues: 1-607

A:Cross-references: EMBL:X17055; NID:g1386; PIDN:CAA34903.1; PTD:g1387
C:Comment: Serum albumin is synthesized in the liver as preproalbumin. It binds cop-
teroid hormones (weak bonds with these hormones promote their transfer across the me-
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-24/Domain: propeptide #status predicted <PRO>
F:25-607/Product: serum albumin #status predicted <MAT>
F:29-201/Domain: serum albumin repeat homology <SAL>
F:220-393/Domain: serum albumin repeat homology <SA2>
F:412-591/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status predicted
F:77-86,99-115,114-125,147-192,191-200,223-269,268-276,288-302,301-312,339-384,383-3
F:263/Binding site: bilirubin (Lys) #status predicted

Query Match 78.48; Score 2432.5; DB 1; Length 607;

Best Local Similarity 75.0%; Pred. No. 2.1e-153;
Matches 437; Conservative 73; Mismatches 72; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFCDLGEENFKALVLIATFAOYLQQCFDEHVKLNVETFEAKTCVADESAAE 60
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 25 DTHKSEAHRFDNLGEENFGIVILARSQYLQQCFDEHVKLNVETFEAKTCVADESAA 84

QY 61 NCKSLHTFLGDKLTCTVATRETYGBMADCCAKQBPERNECFLQHKKDDNPLRLVRPEV 120
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 85 GCDKSLHTFLGDELCAVKTLRTGYMDACCQEKEPNERECFLNHKKDSDLPKL-KPEP 143

QY 121 DVMTAFDHNETELFKLYIARIARRHYFYAPPELLFAFRKYKAFTTECOQAADKAACLLP 180
| : | : | : | : | | | | | | | | | | | : | | | | | : | | | | : | | | | :
Db 144 DTICAEKADEKFFGWGLYEVARRHPTTFAPELLITANKNGVFQEQCAEDKGACLIP 203

QY 181 KLDELDEGGKASSAQRIKLCASLOKFGERAFKAWARLSORFPKAEFAEVSKLVTDLRK 240
| | : | : | | | | | | | | | | | | | | | | : | | | | | : | | | | : | | | | :
Db 204 KIDAMREKVLAASSQRQLRCSIOKFGERALKANSVARLSOKFPKADFDTVIKTVDLRK 263

QY 241 VHEPCCHGDLLCECADRADLAKYCENODISSKLKCECPLELKSHCIAVENDEMPPA 300
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 264 VHKECCHGDLLCECADRADLAKYTCDHQDALSSKLKCECDKPVLKESHCI AEVDKDAPVE 323

QY 301 DLPSIAADFVESKDVCKNYAEAKDNVPFLGMFLXYEARRRHDYSVILLRLRAKTVETLEKC 360
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 324 NLPLETADFAEKVCNKYQZAKOVFLGSFLIEISRHHPEYAVSVILLRLAKEYEATLED 383

QY 361 CAAADPHCYAKVDFEPKLYEEPONLIKONCELFBOLGEYKFQNALLVRYTKPKPOVST 420
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 384 CAKEDPHACYATVFDKLHLVDPEONLIKKNCELPEKHGEYGFQNALIVRYTRKAPQVST 443

QY 421 PTLVEVSRLGVSGKCKHPKAMPADBYLSVVLNQCVLHKPFVSDRVTKCOTES 480
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 444 PTLVEISRLGVSGTKCKAKESERPMPCTDYLSHLNRLCVLHKPFVSEKVTCKOTES 503

QY 481 LVNRRCFSALVEDITYPKEFNAETTFHADICTLSKERQIKQTQALVELYKHKPKAT 540
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 504 LVNRRCFSDLILDITYPKFDEBKFTTFHADICTLTPDEKOIRKQVALVELLHKHPKAT 563

QY 541 KEQLKAVMDDAFVEKCKCKADDKETCFABEGKKLVAASAQAL 583
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 564 DEQLKTNMFVAFVDCCAADDEGCGVLEGFLGVASTQAL 606

RESULT 7
ABTS
serum albumin precursor - rat
N:Alternate names: preproalbumin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 31-May-1979 #sequence_revision 31-May-1979 #text_change 22-Jun-1999
C:Accession: A93872; A92211; A91946; A91940; C45800; I57621; A03233
R:Sargent, T.D.; Yang, M.; Bonner, J.
Proc. Natl. Acad. Sci. U.S.A. 78, 243-246, 1981
A:Title: Nucleotide sequence of cloned rat serum albumin messenger RNA.
A:Reference number: A93872; MUID:81223722; PMID:7017712
A:Accession: A93872
A:Molecule type: mRNA
A:Residues: 1-608 <S>
A:Cross-references: GB:V01222; NID:g55627; PIDN:CAA24532.1; PID:g55628
R:Strauss, A.W.; Bennett, C.D.; Donohue, A.M.; Rodkey, J.A.; Alberts, A.W.
J. Biol. Chem. 252, 6846-6855, 1977
A:Title: Rat liver pre-proalbumin: complete amino acid sequence of the pre-piece. Analysis
of cleavages during protein maturation
A:Reference number: A92211; MUID:77249657; PMID:893447
A>Note: cleavages during protein maturation
A:Accession: A92211
A:Molecule type: protein
A:Residues: 1-38 <S>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 83, 35-48, 1978
A:Title: Amino acid sequences of fragments I and II obtained by cyanogen bromide cleavage
of fragments I and II obtained by cyanogen bromide cleavage
A:Reference number: A91946; MUID:78109429; PMID:564345
A:Accession: A91946
A:Molecule type: protein

A:Residues: 25-222 <SI>
R:Isemura, S.; Ikenaka, T.
J. Biochem. 79, 1183-1196, 1976
A:Title: Fragmentation of rat serum albumin by cyanogen bromide cleavage and the amin
A:Reference number: A91940; MUID:76260153; PMID:956149
A:Accession: A91940
A:Molecule type: protein
A:Residues: 223-288;572-608 <IS>
A>Note: 262-Leu was also found
R:Aoyagi, Y.; Ikenaka, T.; Ichida, F.
Cancer Res. 38, 3483-3486, 1978
A:Title: Copper(II)-binding ability of human alpha-fetoprotein.
A:Reference number: A90758; MUID:79001617; PMID:80265
A:Contents: annotation; copper binding
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid prot
A:Reference number: A45800; MUID:89341406; PMID:2474609
A:Accession: C45800
A>Status: preliminary
A:Molecule type: protein
A:Residues: 166-173 <CAR>
R:Heard, J.
Mol. Cell. Biol. 7, 2425-2434, 1987
A:Title: Determinants of rat albumin promoter tissue specificity analyzed by an impr
A:Reference number: I57621; MUID:87286876; PMID:3475566
A:Accession: I57621
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-5 <RES>
A:Cross-references: GB:M16825; NID:g202828; PIDN:AAA40712.1; PID:g554412
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: carrier protein; duplication; metal binding; plasma
F:1-18/Domain: signal sequence #status experimental <SIG>
F:19-24/Domain: propeptide status experimental <PRO>
F:25-608/Product: serum albumin #status experimental <MAT>
F:229-202/Domain: serum albumin repeat homology <SAI>
F:221-394/Domain: serum albumin repeat homology <SA>
F:413-592/Domain: serum albumin repeat homology <SA3>
F:27/Binding site: copper (His) #status experimental
F:77-86,99-115,114-125,148-193,192-201,224-270,269-277,289-303,302-313,340-385,384-393

Query Match 78.2%; Score 2426; DB 1; Length 608;
Best Local Similarity 73.4%; Pred. No. 5.6e-153;
Matches 428; Conservative 82; Mismatches 73; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFCDLGEENFKALVLIATFAOYLQQCFDEHVKLNVETFEAKTCVADESAAE 60
Db 25 EAHKSIAHRFCDLGEHFGLVLIATFOYLOKCPYESHIKLVQEVTFDAKTCAVDENAE 84

QY 61 NCKSLHTFLGDKLTCTVATRETYGBMADCCAKQBPERNECFLQHKKDDNPLRLVRPEV 120
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 85 NCDSLHTFLGDKLCAIPKLDNKGELADCCAKQBPERNECFLQHKKDDNPLRPQRFEA 144

QY 121 DVMTAFDHNETELFKLYIARIARRHYFYAPELIFFAFRKYKAFTTECOQAADKAACLLP 180
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 145 EAMCTSFOENPTSFLIGHTHEVARRHYPYAPPELLYAEKYNELVITQCCTESDKAACLTTP 204

QY 181 KLDELDEGGKASSAQRIKLCASLOKFGERAFKAWARLSORFPKAEFAEVSKLVTDLRK 240
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 205 KLDPAEKALVAAYRKMSQMRGFERAFKANAVARMRSQRFNFAEFITKLATDVTK 264

QY 241 VHEPCCHGDLLCECADRADLAKYCENODISSKLKCECPLELKSHCIAVENDEMPPA 300
| : | | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 265 INKECCHGDLLCECADRADLAKYCENQATFISSKLOACCDKPVLOKSOCLETEDHNIPA 324

QY 301 DLPSIAADFVESKDVCKNYAEAKDNVPFLGMFLXYEARRHDPYISVILLRLRAKTVETLEKC 360
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 325 DLPSIAADFVEDKEDVCKNYAEAKDNVPFLGYESSRHPDYSVILLRLRAKKEYATLEKC 384

QY 361 CAAADPHCYAKVDFEPKLYEEPONLIKONCELFBOLGEYKFQNALLVRYTKPKPOVST 420
| | | | | : | | | | : | | | | : | | | | : | | | | : | | | | : | | | | :
Db 385 CAREGPPACYGTVLAEPQVLFVEFNKLVKNCIELYELKGYFQNAVLYRYTKAPOYST 444

RESULT 13

FPHU

alpha-fetoprotein precursor [validated] - human

N/Alternate names: AFP; alpha-1-fetoprotein; alpha-fetoglobulin

C/Species: Homo sapiens (man)

C/Date: 19-Feb-1984 #sequence.revision 19-Feb-1984 #text.change 08-Dec-2000

C/Accession: A26624; S37655; A93963; A91497; A23659; A61480; A90624; A90757; A93042;

R/Gibbs, P.E.M.; Zielinski, R.; Boyd, C.; Dugaiczak, A.

Biochemistry 26, 1332-1343, 1987

A/Title: Structure, polymorphism, and novel repeated DNA elements revealed by a comp

A/Reference number: A26624; MUID:87185438; PMID:2436661

A/Accession: A26624

A/Molecule type: DNA

A/Residues: 1-609 <GIB>

A/Cross-references: GB:M16110; NID:g773678; PIDN:AAB58754.1; PID:g178236

R/McVey, J.H.; Michaelides, K.; Hansen, L.P.; Ferguson-Smith, M.; Tilghman, S.; Krum

Hum. Mol. Genet. 2, 379-384, 1993

A/Title: A G->A substitution in an HNF I binding site in the human alpha-fetoprotein

A/Reference number: S37655; MUID:93278385; PMID:7684942

A/Accession: S37655

A/Molecule type: DNA

A/Residues: 1-28 <MCV>

A/Cross-references: EMBL:Z19532; NID:z28527; PIDN:CAA79592.1; PID:g28528

A/Note: The authors translated the codon TAT for residue 26 as Thr

R/Morinaga, T.; Sakai, M.; Wegmann, T.G.; Tamaoki, T.

Proc. Natl. Acad. Sci. U.S.A. 80, 4604-4608, 1983

A/Title: Primary structures of human alpha-fetoprotein and its mRNA.

A/Reference number: A93961; MUID:83273664; PMID:6192439

A/Accession: A93961

A/Molecule type: mRNA

A/Residues: 1-609 <MOR>

A/Cross-references: GB:J00077; NID:g311348; PIDN:CAA24758.1; PID:g31351

R/Beattie, W.G.; Dugaiczak, A.

Gene 20, 415-422, 1982

A/Title: Structure and evolution of human alpha-fetoprotein deduced from partial seq

A/Reference number: A91497; MUID:83158778; PMID:6187626

A/Accession: A91497

A/Molecule type: mRNA

A/Residues: 429-556 <BEA>

A/Cross-references: GB:J00076

R/Pucci, P.; Sciallano, R.; Malorni, A.; Marino, G.; Tecce, M.F.; Ceccarini, C.; Ter

Biochemistry 30, 5061-5066, 1991

A/Title: Human alpha-fetoprotein primary structure: a mass spectrometric study.

A/Reference number: A23659; MUID:91242409; PMID:1709810

A/Accession: A23659

A/Molecule type: protein

A/Residues: 19-45;60-97;102-107;122-184;187-249;255-489;507-609 <PUC>

R/Tecce, M.F.; Terrana, B.; Giuliani, M.M.; Ceccarini, C.

J. Nucl. Med. Allied Sci. 34, 213-216, 1990

A/Title: Characterization of in vitro expressed human alpha-fetoprotein as highly re

A/Reference number: A61480; MUID:91225826; PMID:1709209

A/Accession: A61480

A/Molecule type: protein

```

Db      141  VPPEVTSCEAYEEDRETFMKNF IYEIARRHPFLYAPTILLWAARYDKILPSCCKAEAYE 200
QY      177  CLFLKLELDRBGKASSAKQELKQKASLQKFGFBRFAKAWARLSQRPKAEFAEYSLKVT 236
Db      201  CFQTKAATVTKELRESSLLNQHACAYMKNFGTTFQAITVTFKLSQKTKVNFTEIQKLWL 260
QY      237  DLTKVHTECHGDGDLLECADDRLADLAKYICENODSISLKKECCEKPLLEKSHCIAEVND 296
Db      261  DVARVHEHCRCRGVLDCCIQDEKIMSYGCSQDQTSNRTICECKLTLTLERGQCIIHAEND 320
QY      297  EMPADPLSLAAADFBSKVCKNAEAKDVFGLMFLEYEARHPDYSVLLILRLAKTYETT 356
Db      321  EKPEGLSPNLNFIQDRDNFNFSSGEKNIFLASFVHEYSRRHPQLAVSVILVAKGYOEL 380
QY      357  LEKCCAAADPHCYAKVDEPKPLVIBEPONLIKQNCSELFEOIGYKFNQNALIVRTTKVP 416
Db      381  LEKQFOTENPLECDQKGBEEELQRTIQESQALAKRSCGLFQKLGEBYLLQNAFLVATTKAP 440
QY      417  QVSTPTLVSVSNLKGKVGCKCKHPFAKMPCAEDYTLVSNLNLQVLVHEKTPVSDRVTKC 476
Db      441  QLTSSSELMATRKMAATAATCCQLSEDKLLAGGGAADIIIGHLCIRHEMTVPNFGVQC 500
QY      477  CTBSLVNRRPCTSALEVDYVVPKERNATFTTHADICTLSEKERQTKKATALVELNKH 536
Db      501  CTSSYANRRPCFSLSLWDEVTVPFATSDQKTFPHKDLQQAQGVALQPMKQEFILNLVKQ 560
QY      537  PKATKQLKAWMDDFAAVEKCKCADDKCTCAEBSGKKLVAAASAALGL 585
Db      561  POITPEQLAVIATDGLLEKKCCQOGQGVCFVAFBGGKLLISKTRALGV 609

RESULT 14
PGCG
alpha-fetoprotein precursor - gorilla
C:Species: Gorilla gorilla (gorilla)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 22-Jun-1999
C:Accession: A37970
R:Ryan, S.C.; Zielinski, R.; Dugaiczky, A.
Genomics 9, 60-72, 1991
A:title: Structure of the gorilla alpha-fetoprotein gene and the divergence of primat
A:Reference number: A37970; MUID:91169517; PMID:1706310
A:Accession: A37970
A:Molecule type: DNA
A:Residues: 1-609 <R>A>
A:Cross-references: GB:M38272; NID:g817963; PIDN:AAA73520.1; PID:gi77041
C:Genetics:
A:Map position: 4q11-12
A:Introns: 29/1; 46/2; 90/3; 161/2; 205/3; 238/2; 281/3; 353/2; 397/3; 430/2; 476/3;
C:Superfamily: serum albumin; serum albumin repeat homology
C:Keywords: embryo; fetus; globulin; glycoprotein; metal binding; plasma
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-609/Product: alpha-fetoprotein #status predicted <M>A>
F:29-202/Domain: serum albumin repeat homology <S>A1>
F:221-394/Domain: serum albumin repeat homology <S>A2>
F:413-592/Domain: serum albumin repeat homology <S>A3>
F:22/binding site: copper (His) #status predicted
F:99-114,113-124,148-193,192-201,224-270,269-277,289-303,302-313,384-393,416-462,461-
F:249/Binding site: bilirubin (Lys) #status predicted
F:351/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match      40.0%; Score 1242.5; DB 1; Length 609;
Best Local Similarity 39.6%; Pred. No. 1.le-74;
Matches 233; Conservative 117; Mismatches 232; Indels 7; Gaps 3;

QY 3 HKSE-----VAHRFKDLGEEFKALVLIAFALQCCPFEDHVKLVNEVTEFAKTCVADE 57
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 22 HRNEYGTASILDSYQCTREISLADLATITTFAGVQGYTKVSMWKDALTALEKPTGDE 81
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

QY 58 SAENCOKSLHTLFGDKLQCVATLRETYGEMADCAQKQEPERNECFLOHQKDDNP-NLPLRV 116
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|
Db 82 QSACCLNQLPAFLELCEHEKLEIKYG-LSQCCSQSEEGRHNCFLAHKKPFPASPILEQ 140
   |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:| |:|

QY 117 RPEYVMCTAFHDNEETFLKYLEIYLRHRPHYFVAPELLFFAKRYKAAFTTCCQAAKAA 176

```

Search completed: August 31, 2003, 16:34:37
Job time : 45 secs

Search completed: August 31, 2003, 16:34:37
Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 25 Seconds
(without alignments)
1100.425 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103

Sequence: 1 DAKSVAHRFDLGEENFK.....TCFABEGKKLVASQAALGL 585

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3103	100.0	609	1	ALBU_HUMAN
2	2942	94.8	600	1	ALBU_MACMU
3	2620	84.4	608	1	ALBU_FELCA
4	2562	82.6	608	1	ALBU_CANFA
5	2475.5	79.8	607	1	ALBU_HORSE
6	2450.5	79.0	607	1	ALBU_BOVIN
7	2446	78.8	608	1	ALBU_SHEEP
8	2432.5	78.4	607	1	ALBU_RAT
9	2426	78.2	608	1	ALBU_PIG
10	2411.5	77.7	605	1	ALBU_MERUN
11	2387	76.9	609	1	ALBU_MOUSE
12	2378	76.6	608	1	ALBU_CHICK
13	1557.5	50.2	615	1	ALBU_CHICK
14	1253.5	40.4	609	1	FETA_PANTR
15	1249.5	40.3	609	1	FETA_HUMAN
16	1242.5	40.0	609	1	FETA_GORGO
17	1205	38.8	607	1	ALB2_XENLA
18	1200	38.7	609	1	FETA_HORSE
19	1164.5	37.5	606	1	ALB1_XENLA
20	1084	34.9	605	1	FETA_MOUSE
21	1067	34.4	611	1	FETA_RAT
22	1055	34.0	599	1	AFAM_HUMAN
23	944	30.4	611	1	AFAM_MOUSE
24	928	29.9	608	1	AFAM_RAT
25	747.5	24.1	608	1	ALB1_SALSA
26	742.5	23.9	608	1	ALB2_SALSA
27	699	22.5	382	1	ALBU_RANCA
28	440.5	14.2	1423	1	ALBU_PETMA
29	386	12.4	474	1	VTDB_HUMAN
30	381	12.3	476	1	VTDB_RAT
31	378	12.2	476	1	VTDB_RABIT
32	372	12.0	472	1	VTDB_MOUSE
33	151.5	4.9	1605	1	RRE1_MOUSE

34 144.5 4.7 8797 1 SNE1_HUMAN
35 138.5 4.5 1410 1 RRE1_HUMAN
36 133.5 4.3 1391 1 MST2_DROHY
37 132.5 4.3 2230 1 GOG4_HUMAN
38 129.5 4.2 1972 1 MYHB_HUMAN
39 129 4.2 3210 1 CENF_HUMAN
40 128 4.1 1005 1 RA50_MERJA
41 126.5 4.1 1972 1 MYHB_RABIT
42 126 4.1 1189 1 SMC2_CHICK
43 126 4.1 3259 1 GIAN_HUMAN
44 125 4.0 3038 1 TRIO_HUMAN
45 124.5 4.0 1790 1 USO1_YEAST

ALIGNMENTS

RESULT 1
ALBU_HUMAN
ID ALBU_HUMAN STANDARD; PRT; 609 AA.
AC P02768; O95574; Q13140; Q9P157; Q9P17; Q9UHS3; Q9UJZ0;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86196112; PubMed=3009475;
RA Minghetti P.P., Rufner D.E., Kuang W.J., Dennison O.E., Hawkins J.W.,
RA Beattie W.G., Dugaiczky A.;
RT "Molecular structure of the human albumin gene is revealed by
RT nucleotide sequence within g11-22 of chromosome 4.";
RL J. Biol. Chem. 261:6747-6757(1986).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT LYS-420.
RX MEDLINE=8201882; PubMed=6171778;
RA Lawn R.M., Adelman J., Bock S.C., Franke A.E., Houck C.M.,
RA Najarian R.C., Seeburg P.H., Wilson K.L.;
RT "The sequence of human serum albumin cDNA and its expression in E.
RT coli.";
RL Nucleic Acids Res. 9:6103-6114(1981).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANT GLY-121.
RX MEDLINE=82105994; PubMed=6275391;
RA Dugaiczky A., Law S.W., Dennison O.E.;
RT "Nucleotide sequence and the encoded amino acids of human serum
RT albumin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:71-75(1982).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RT "Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases."
RN [5]
RP SEQUENCE FROM N.A. (PRO0903/PRO1708/PRO2044/PRO2619/PRO2675).
RX TISSUE=Fetal liver;
RT "Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases."
RN [6]
RP SEQUENCE FROM N.A., AND VARIANT HIROSHIMA-1 LYS-378.
RX Huang M.C., Wu H.T.;
RT "The cDNA sequences of human serum albumin.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.

RC TISSUE=Liver, and Skeletal muscle;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [8]
RP SEQUENCE OF 25-609.
RX MEDLINE=76187907; PubMed=1225573;
RA Meloun B., Moravsek L., Kostka V.;
RT "Complete amino acid sequence of human serum albumin.";
RL FEBS Lett. 58:134-137(1975).
RN [9]
RP SEQUENCE OF 25-609.
RX Brown J.R., Shockley P., Behrens P.Q.;
RL (In) Bing D.H. (eds.);
RL "The chemistry and physiology of the human plasma proteins, pp.23-40,
RL Pergamon Press, New York (1979).
RN [10]
RP SEQUENCE OF 1-455 FROM N.A.
RC TISSUE=Liver;
RA Menaya J., Parrilla R., Ayuso M.S.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [11]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=86140099; PubMed=2419329;
RA Urano Y., Watanabe K., Sakai M., Tamaoki T.;
RT "The human albumin gene. Characterization of the 5' and 3' flanking
RT regions and the polymorphic gene transcripts.";
RL J. Biol. Chem. 261:3244-3251(1986).
RN [12]
RP SEQUENCE OF 222-229.
RX MEDLINE=76257808; PubMed=955075;
RA Walker J.E.;
RT "Lysine residue 199 of human serum albumin is modified by
RT acetylsalicylic acid.";
RL FEBS Lett. 66:173-175(1976).
RN [13]
RP SEQUENCE OF 25-44 AND 480-499.
RC TISSUE=Heart;
RX MEDLINE=95203287; PubMed=7895732;
RA Corbett J.M., Wheeler C.H., Baker C.S., Yacoub M.H., Dunn M.J.;
RT "The human myocardial two-dimensional gel protein database: update
RT 1994.";
RL Electrophoresis 15:1459-1465(1994).
RN [14]
RP DISULFIDE BONDS.
RA Sabar M.A., Stockbauer P., Moravsek L., Meloun B.;
RT "Disulfide bonds in human serum albumin.";
RL Collect. Czech. Chem. Commun. 42:564-579(1977).
RN [15]
RP BILIRUBIN-BINDING SITE.
RX MEDLINE=78186630; PubMed=656055;
RA Jacobsen C.;
RT "Lysine residue 240 of human serum albumin is involved in high-
RT affinity binding of bilirubin.";
RL Biochem. J. 171:453-459(1978).

RN [16]
RX VARIANT CANTERBURY ASN-337.
RA MEDLINE=87157744; PubMed=3828358;
RA Brennan S.O., Herbert P.;
RT "Albumin Canterbury (313 Iys->Asn). A point mutation in the second
RT domain of serum albumin.";
RL Biochim. Biophys. Acta 912:191-197(1987).
RN [17]
RP VARIANTS NAG-2 AND NAG-3.
RX MEDLINE=88068523; PubMed=3479777;
RA Takahashi N., Takahashi Y., Isobe T., Putnam F.W., Fujita M.,
RA Satoh C., Neel J.V.;
RT "Amino acid substitutions in inherited albumin variants from
RT Amerindian and Japanese populations.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8001-8005(1987).
RN [18]
RP VARIANTS NAG-1; HIR-1; HIR-2 AND TOCHIGI.
RX MEDLINE=89345611; PubMed=2762316;
RA Arai K., Madison J., Huss K., Ishioka N., Satoh C., Fujita M.,
RA Neel J.V., Sakurabayashi I., Putnam F.W.;
RT "Point substitutions in Japanese alloalbumins.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:6092-6096(1989).
RN [19]
RP VARIANTS MANAUS; OSAKA; NAGOYA; FUKUOKA; HONOLULU AND NEW-GUINEA.
RX MEDLINE=90115905; PubMed=2404284;
RA Arai K., Madison J., Shimizu A., Putnam F.W.;
RT "Point substitutions in albumin genetic variants from Asia.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:497-501(1990).
RN [20]
RP DESCRIPTION OF VARIANT REDHILL.
RX MEDLINE=90115852; PubMed=2104980;
RA Brennan S.O., Myles T., Peach R.J., Donaldson D., George P.M.;
RT "Albumin Redhill (-1 Arg, 320 Ala->Thr): a glycoprotein variant of
RT human serum albumin whose precursor has an aberrant signal peptidase
RT cleavage site.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:26-30(1990).
RN [21]
RP VARIANTS TORINO LYS-84; VARESE HIS-23 AND VIBO VALENTIA LYS-106.
RX MEDLINE=91062352; PubMed=2247440;
RA Galliano M., Minchiotti L., Porta F., Rossi A., Ferri G., Madison J.,
RA Watkins S., Putnam F.W.;
RT "Mutations in genetic variants of human serum albumin found in
RT Italy.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8721-8725(1990).
RN [22]
RP VARIANT VENEZIA.
RX MEDLINE=91296740; PubMed=2068071;
RA Watkins S., Madison J., Davis E., Sakamoto Y., Galliano M.,
RA Minchiotti L., Putnam F.W.;
RT "A donor splice mutation and a single-base deletion produce two
RT carboxyl-terminal variants of human serum albumin.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:5959-5963(1991).
RN [23]
RP VARIANTS IOWA CITY-2 VAL-25; IOWA CITY-1 VAL-389; KOMAGOME-3 HIS-23;
RX KOMAGOME-2 ARG-152 AND KOMAGOME-1 GLU-396.
RA MEDLINE=92052189; PubMed=1946412;
RA Madison J., Arai K., Feld R.D., Kyle R.A., Watkins S., Davis E.,
RA Matsuda Y.-I., Amaki I., Putnam F.W.;
RT "Genetic variants of serum albumin in Americans and Japanese.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9853-9857(1991).
RN [24]
RP VARIANT CASEBROOK ASN-518.
RX MEDLINE=91316157; PubMed=1859851;
RA Peach R.J., Brennan S.O.;
RT "Structural characterization of a glycoprotein variant of human serum
RT albumin: albumin Casebrook (494 Asp->Asn).";
RL Biochim. Biophys. Acta 1097:49-54(1991).
RN [25]
RP VARIANTS SONDRIO LYS-357 AND PARIS-2 ASN-587.
RX MEDLINE=92190239; PubMed=1347703;
RA Minchiotti L., Galliano M., Stoppini M., Ferri G., Crespeau H.,
RA Rochu D., Porta F.;
RT "Two alloalbumins with identical electrophoretic mobility are produced

```

Query Match      100.0%; Score 3103; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 1.7e-194;
Matches 585; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60
DB 25 DAHSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 84
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 120
DB 85 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 144
QY 121 DVMTAFPHNEETFLKKYLYEIARRHPHYFAPPELLFFAKRYKAFTCCQAADKAACLLP 180
DB 145 DVMTAFPHNEETFLKKYLYEIARRHPHYFAPPELLFFAKRYKAFTCCQAADKAACLLP 204
QY 181 KLDELDEGKASSAQRLKCSLQKFGERAFAKAWAVARLSQFFKAEFAEVSCLVTDLTK 240
DB 205 KLDELDEGKASSAQRLKCSLQKFGERAFAKAWAVARLSQFFKAEFAEVSCLVTDLTK 264
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 265 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 324
QY 301 DLPSLAADFVESKDVCKNYAKADVFLGMFLYERARRHPDYSVILLRLAKTYETTLK 360
DB 325 DLPSLAADFVESKDVCKNYAKADVFLGMFLYERARRHPDYSVILLRLAKTYETTLK 384
QY 361 CAADAPHECYAKVDFEKLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 420
DB 385 CAADAPHECYAKVDFEKLVEEPQNLKONCELFQELGEYKFNALLVRYTKVPQVST 444
QY 421 PTLVEVSRNLGVSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRYTKCCTES 480
DB 445 PTLVEVSRNLGVSKCKHPEAKMPCAEADYLSVLNQLCVLHEKTPVSDRYTKCCTES 504
QY 481 LVNRRCFSALEVDYETVYKFEFNAETTFHADICTLSEKEROIKQIALVELYKHKPKAT 540
DB 505 LVNRRCFSALEVDYETVYKFEFNAETTFHADICTLSEKEROIKQIALVELYKHKPKAT 564
QY 541 KEQLKAVMDODFAFVEKCKADKCTCFABEGKKLVAASQAALGL 585
DB 565 KEQLKAVMDODFAFVEKCKADKCTCFABEGKKLVAASQAALGL 609

RESULT 2
ALBU_MACMU STANDARD; PRG; 600 AA.
AC Q28522;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor (Fragment).
GN ALB.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93211971; PubMed=8460152;
RA Watkins S.A., Sakamoto Y., Madison J.M., Davis E.M., Smith D.G.,
RA Dwulet J., Putnam F.W.;
RT "cDNA and protein sequence of polymorphic macaque albumins that differ
in bilirubin binding.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:2409-2413(1993).
CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
hormones, bilirubin and drugs. Its main function is the regulation
of the colloidal osmotic pressure of blood.
CC -!- SUBCELLULAR LOCATION: Secreted.

```

```

CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: BELONGS TO THE ALB/AFB/VDB FAMILY.
CC -!- SIMILARITY: Contains 3 albumin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M90463; AAA36906.1;
DR PIR: A47391; A47391.
DR HSSP: P02768; 1E7B.
DR InterPro: IPR000264; Serum_albumin.
DR Pfam: PF00273; transport_prot; 3.
DR PRINTS: PR00802; SERUMALBUMIN.
DR PRODOM: PD002486; Serum_albumin; 1.
DR SMART: SM00103; ALBUMIN; 3.
DR PROSITE: PS00212; ALBUMIN; 3.
KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
FT SIGNAL <1 10 10 BY SIMILARITY.
FT PROPEP 11 16 BY SIMILARITY.
FT CHAIN 17 600 SERUM ALBUMIN.
FT DOMAIN 17 197 ALBUMIN 1.
FT DOMAIN 204 389 ALBUMIN 2.
FT DOMAIN 396 587 ALBUMIN 3.
FT METAL 19 19 COPPER (BY SIMILARITY).
FT BINDING 256 256 BILIRUBIN (POTENTIAL).
FT DISULFID 69 78 BY SIMILARITY.
FT DISULFID 91 107 BY SIMILARITY.
FT DISULFID 106 117 BY SIMILARITY.
FT DISULFID 140 185 BY SIMILARITY.
FT DISULFID 184 193 BY SIMILARITY.
FT DISULFID 216 262 BY SIMILARITY.
FT DISULFID 261 269 BY SIMILARITY.
FT DISULFID 281 295 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 332 377 BY SIMILARITY.
FT DISULFID 376 385 BY SIMILARITY.
FT DISULFID 408 454 BY SIMILARITY.
FT DISULFID 453 464 BY SIMILARITY.
FT DISULFID 477 493 BY SIMILARITY.
FT DISULFID 492 503 BY SIMILARITY.
FT DISULFID 530 575 BY SIMILARITY.
FT DISULFID 574 583 BY SIMILARITY.
SQ SEQUENCE 600 AA; 67880 MW; E45C871A670E740B CRC64;

Query Match      94.8%; Score 2942; DB 1; Length 600;
Best Local Similarity 93.5%; Pred. No. 4.8e-184;
Matches 545; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60
DB 17 DTHKSEVAHRFDLGEENFKGLVLVAFSOLQCPEDHVKLVNEVTEFAKTCVADSEAE 76
QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 120
DB 77 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLRLVPRV 136
QY 121 DVMTAFPHNEETFLKKYLYEIARRHPHYFAPPELLFFAKRYKAFTCCQAADKAACLLP 180
DB 137 DVMTAFPHNEATFLKKYLYEIARRHPHYFAPPELLFFAKRYKAFTCCQAADKAACLLP 196
QY 181 KLDELDEGKASSAQRLKCSLQKFGERAFAKAWAVARLSQFFKAEFAEVSCLVTDLTK 240
DB 197 KLDELDEGKASSAQRLKCSLQKFGERAFAKAWAVARLSQFFKAEFAEVSCLVTDLTK 256
QY 241 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 300
DB 257 VHTCECHGDLLECCADDRADLAKYICENQDSISSKLEKCEKPLLEKSHCIAEVENDEMPA 316

```

CC	EMBL; X84842; CAA59279.1; -.
DR	PIR; JC4660; S57632.
DR	HSSP; P02768; IE7B.
DR	InterPro; IPR000264; Serum_albumin.
DR	Pfam; PF00273; transport_prot; 3.
DR	PRINTS; PRO0802; SERUMALBUMIN.
DR	ProDom; PD002486; Serum_albumin; 1.
DR	SMART; SM00103; ALBUMIN; 3.
DR	PROSITE; PS00212; ALBUMIN; 3.
KW	Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.
FT	SIGNAL 1 18
FT	PROPEP 19 24
FT	CHAIN 25 608
FT	DOMAIN 25 205
FT	DOMAIN 212 397

Query Match	84.4%;	Score 2620;	DB 1;	Length 608;
Best Local Similarity	82.0%;	Pred. No. 4e-163;		
Matches 478;	Conservative 52;	Mismatches 53;	Indels 0;	Gaps 0;
QY	1	DAKSEVAHFEDLGEENFKALVILIAFAOYLQOCPEDHVLVNEVTEFAKTCVADSEAE	60	
Db	25	EAHOSEIAHFNLDGEHFHGLVIAFVSQYLQOCPEDHVLVNEVTEFAKGCVADSEA	84	
QY	61	NCDSLHTLFGDKLCTVATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPRVPRV	120	
Db	85	NCEKSLHELLGDKLCTVASLRDKYEMADCCCKEPERNECFLOHKDDNPGFQQLTPPA	144	
QY	121	DVMTAFPHDNEETFLKKLYEIAIRRHYPFYVAPELLPFAKRYKAAFTSCCOAADAACLA	180	
Db	145	DAMCTAPHENQRFGLKLYEIAIRRHYPFYVAPELLYIAYEYTGVTTCCEAADAACLA	204	
QY	181	KILDERDEGKASSAKORLKCSLOKFGERAFKAWAVARLSORFPAKFAFVSKVLVTDL	240	
Db	205	KVDALREKVLASSAKERLKCASLOKFGERAFKAWAVARLSORFPAKFAFVSKVLVTD	264	
QY	241	VHTECCHGDLLECAADRADLAKYICENQDSTISKLKECCERPLLEKSHCIAFVEND	300	
Db	265	IHKCECHGDLLECAADRADLAKYICENQDSTISKLKECCGKPVLEKSHCISEVER	324	
QY	301	DLPSLAADFVSKDCKVKNYBAKDVIFGMFLFYEARHGDYSVLLRLAKATTETTL	360	
Db	325	DLPLAVDFVEDKVCNKYQAKDVLFTGLTYEYSRRHPEYSVLLRLAKAYEATLE	384	
QY	361	CAAADPHECYAKVDFDEFPLVEEQNLIKONCELPFLGEYKFQNALLYRYTKKVP	420	
Db	385	CATDDPACVYAHVDFDFKPLVEEPHNLVKTNCLEFPEKGEYGFQNALLYRYTKK	444	
QY	421	PTLVEVSRLNGKVGSKCKHPKARMPCAEDYLSVYNOLCVLHEKTPVSDRYTKCT	480	
Db	445	PTLVEVSRLNGKVGSKCKTHPEARLSCAEDYLSVYNLPLCVLHEKTPVSERYTK	504	
QY	481	LVNRRPCFSALEVDEYTPYKFEFAETFTFHADICTLSEKERQIKQOTALVELYKH	540	
Db	505	LVNRRPCFSALEVDEYTPYKFSATFTFHADICTLSEPAEKQIKQOTALVELYKH	564	

RESULT 4	
ALBU CANFA	
ID	ALBU CANFA
STANDARD;	
AC	A49822; 077705; Q9TSZ4;
DT	01-OCT-1996 (Rel. 34, Created)
DT	15-OCT-2001 (Rel. 40, Last sequence update)
DT	16-FEB-2003 (Rel. 41, Last annotation update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)

DE Serum albumin precursor (Allergen Can f 3).

GN ALB.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=Beagle; TISSUE=Liver;

RA Hilger C.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RC MEDLINE=20148667; PubMed=10669848;

RA Pandjaitan B., Swoboda I., Brandesky-Pichler F., Rumpold H.,

RA Valenta R., Spitzauer S.;

RT "Escherichia coli expression and purification of recombinant dog

RT albumin, a cross-reactive animal allergen.";

RL J. Allergy Clin. Immunol. 105:279-285(2000).

RN [3]

RN SEQUENCE OF 25-48.

RP MEDLINE=75011422; PubMed=4414013;

RA Dixon J.W., Sarkar B.;

RT "Isolation, amino acid sequence and copper(II)-binding properties of

RT peptide (1-24) of dog serum albumin.";

RL J. Biol. Chem. 249:5872-5877(1974).

RN [4]

RN SEQUENCE OF 25-38.

RC TISSUE=Heart;

RC MEDLINE=98163340; PubMed=9504812;

RA Dunn M.J., Corbett J.M., Wheeler C.H.;

RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of

RT dog heart proteins.";

RL Electrophoresis 18:2795-2802(1997).

RN [5]

RN SEQUENCE OF 215-478 FROM N.A.

RC TISSUE=Salivary gland;

RC MEDLINE=94201492; PubMed=7512102;

RA Spitzauer S., Schweizer C., Sperr W.R., Pandjaitan B., Valent P.,

RA Muehl S., Ebner C., Scheiner O., Kraft D., Rumpold H.;

RT "Molecular characterization of dog albumin as a cross-reactive

RT allergen.";

RL J. Allergy Clin. Immunol. 93:614-627(1994).

CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Plasma.

CC -1- SIMILARITY: BELONGS TO THE ALB/APP/YDB FAMILY.

CC -1- SIMILARITY: Contains 3 albumin domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AJ133489; CAB64867.1; -

DR EMBL; Y17737; CA76841.1; -

DR EMBL; S72946; AAB30434.1; -

DR HSP; P02768; 1E7B.

DR HSC-2DPAGE; P49822; DOG.

DR InterPro: IPR000264; Serum_albumin.

DR Pfam; PF00273; transprot_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR PRODOM; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 3.

KW Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen.

FT SIGNAL 1 18

FT PROPEP 19 24

FT CHAIN 25 608

FT DOMAIN 25 205

FT DOMAIN 212 397

FT DOMAIN 404 595

FT METAL 27 27

FT DISULFID 77 86

FT DISULFID 99 115

FT DISULFID 114 125

FT DISULFID 148 193

FT DISULFID 192 201

FT DISULFID 224 270

FT DISULFID 269 277

FT DISULFID 289 303

FT DISULFID 302 313

FT DISULFID 340 385

FT DISULFID 384 393

FT DISULFID 416 462

FT DISULFID 461 472

FT DISULFID 485 501

FT DISULFID 500 511

FT DISULFID 538 583

FT DISULFID 582 591

FT CONFLICT 1 26

FT CONFLICT 146 146

FT CONFLICT 206 206

FT CONFLICT 349 349

FT CONFLICT 359 359

FT CONFLICT 448 448

FT CONFLICT 474 474

SQ SEQUENCE 608 AA; 68606 MW; 3CF1C8F7DD8FC06 CRC64;

Query Match 82.6%; Score 2562; DB 1; Length 608;

Best Local Similarity 79.8%; Pred. No. 2.4e-159;

Matches 465; Conservative 57; Mismatches 61; Indels 0; Gaps 0;

QY 1 DAHSEVAHRFDKIGENFKALVLIAPAYLQCCFEDHVKLVNVEYFAKTCVADESA 60

DB 25 EAYKSEIAHRYNDLGEHFRGLVAFSQTQCCFEDHVKLVNVEYFAKTCVADESA 84

QY 61 NCKSKLHTFLGDKLCTVATLETYGENADCCAKQEPERNECFLOHKDDNPNLPLRVPV 120

DB 85 NCKSKLHTFLGDKLCTVATLETYGENADCCAKQEPERNECFLOHKDDNPNLPLRVPV 144

QY 121 DVMTAFADHNEFTFLKLYLVIARHPYFYAPPELLFAKRYKAAFTCCQADKAACILP 180

DB 145 DALCAAFQDNEQLFLGKLYLVIARHPYFYAPPELLYAAQYKGVFAECQADKAACILP 204

QY 181 KIDELDEGKASSAKORLKCSAQKGFGRFAKAWARLSQRFKAEFAVSKLVDTLTK 240

DB 205 KIDELDEGKASSAKORLKCSAQKGFGRFAKAWARLSQRFKAEFAVSKLVDTLTK 264

QY 241 VHTCCGHDLLCADDADLAKYICENQDSISSKKECCERPLEKSHCHIAFVNDMPA 300

DB 265 VHKCCGHDLLCADDADLAKYICENQDSISSKKECCERPLEKSHCHIAFVNDMPA 324

QY 301 DLPSLAADFVSKCKVNAKADKVFGLMFLYVYARRHPDYSVVLLRLAKTVEETLEK 360

DB 325 DLPSLAADFVSKCKVNAKADKVFGLMFLYVYARRHPDYSVVLLRLAKTVEETLEK 384

QY 361 CAADPFECYAKVDFEFPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKPVQST 420

DB 385 CATDDPTCYAKVDFEFPLVEEPQNLIKQNCLEFQELGEYKFNALLVRYTKKPVQST 444

QY 421 PTLVSVRNLGKVGSKCKHPKAKMPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480

DB 445 PTLVSVRNLGKVGSKCKHPKAKMPCADYLSVVLNQLCVLHEKTPVSDRVTKCCTES 504

QY 481 LVNRPFCFSALEVDYETVYKPEFNAETTFHADICTLSEKQIKKOTALVELVKKPKAT 540

DB 505 LVNRPFCFSALEVDYETVYKPEFNAETTFHADICTLSEKQIKKOTALVELVKKPKAT 564

RA Barry T., Power S., Gannon F.;
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 RP [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Hilger C., Grigioni F., de Beaufort C., Michel G., Hentges F.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANT THR-214.
 RA Wu H.T., Huang M.C.;
 RL "The complete cDNA sequence of bovine serum albumin";
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1-32.
 RX MEDLINE=80024278; PubMed=488109;
 RA McGillivray R.T.A., Chung D.W., Davie E.W.;
 RT "Biosynthesis of bovine plasma proteins in a cell-free system. Amino-terminal sequence of prealbumin";
 RL Eur. J. Biochem. 98:477-485(1979).
 RN [6]
 RP SEQUENCE OF 25-424 AND 429-607, AND VARIANT THR-214.
 RA Brown J.R.;
 RL "Structure of bovine serum albumin";
 RT Fed. Proc. 34:591-594(1975).
 RN [7]
 RP REVISIONS TO 190-195.
 RA Brown J.R.;
 RL Submitted (APR-1975) to the PIR data bank.
 RN [8]
 RP SEQUENCE OF 402-433.
 RX MEDLINE=82023364; PubMed=7283978;
 RA Reed R.G., Putnam F.W., Peters T. Jr.;
 RT "Sequence of residues 400-403 of bovine serum albumin";
 RL Biochem. J. 191:867-868(1980).
 RN [9]
 RP SEQUENCE OF 19-28.
 RX MEDLINE=77134075; PubMed=843354;
 RA Patterson J.B., Geller D.M.;
 RT "Bovine microsomal albumin: amino terminal sequence of bovine proalbumin";
 RL Biochem. Biophys. Res. Commun. 74:1220-1226(1977).
 RN [10]
 RP SEQUENCE, AND REVISIONS TO 118-119 AND 180.
 RX MEDLINE=91083649; PubMed=2260975;
 RA Hirayama K., Akashi S., Furuya M., Fukuhara K.-I.;
 RT "Rapid confirmation and revision of the primary structure of bovine serum albumin by ESI/MS and Frit-FAB LC/MS";
 RL Biochem. Biophys. Res. Commun. 173:639-646(1990).
 RN [11]
 RP SEQUENCE OF 25-41.
 RX MEDLINE=88267456; PubMed=3389500;
 RA Hsieh J.C., Lin F.P., Tam M.F.;
 RT "Electroblotting onto glass-fiber filter from an analytical isoelectrofocusing gel: a preparative method for isolating proteins for N-terminal microsequencing";
 RL Anal. Biochem. 170:1-8(1988).
 RN [12]
 RP SEQUENCE OF 437-451.
 RA Vilbois F.;
 RL Submitted (AUG-1998) to the SWISS-PROT data bank.
 RN [13]
 RP DISULFIDE BONDS.
 RA Brown J.R.;
 RL "Structure of serum albumin: disulfide bridges";
 RT Fed. Proc. 33:1389-1389(1974).
 CC -1- FUNCTION: Serum albumin, the main protein of plasma, has a good binding capacity for water, Ca(2+), Na(+), K(+), fatty acids, hormones, bilirubin and drugs. Its main function is the regulation of the colloidal osmotic pressure of blood.
 CC CC
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC
 CC -1- TISSUE SPECIFICITY: Plasma.
 CC
 CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
 CC
 CC -1- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M73993; AAA51411.1; -;
 DR EMBL; X58989; CAA41735.1; -;
 DR EMBL; Y17789; CAA76847.1; -;
 DR EMBL; AF542068; AAN17824.1; -;
 DR HSSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR PRODOM; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper; Allergen; Polymorphism.
 KW SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 607
 FT DOMAIN 25 204
 FT DOMAIN 211 396
 FT DOMAIN 403 594
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 147 192
 FT DISULFID 191 200
 FT DISULFID 223 269
 FT DISULFID 268 276
 FT DISULFID 288 302
 FT DISULFID 301 312
 FT DISULFID 339 384
 FT DISULFID 383 392
 FT DISULFID 415 461
 FT DISULFID 460 471
 FT DISULFID 484 500
 FT DISULFID 499 510
 FT DISULFID 537 582
 FT DISULFID 581 590
 FT VARIANT 214 214
 FT CONFLICT 302 302
 FT CONFLICT 304 305
 FT CONFLICT 324 324
 FT CONFLICT 394 395
 FT CONFLICT 437 437
 FT CONFLICT 493 494
 SQ SEQUENCE 607 AA; 69293 MW; 39167DFE768585D4 CRC64;
 A -> T.
 C -> K (IN REF. 6).
 KP -> PC (IN REF. 6).
 N -> D (IN REF. 6).
 ST -> TS (IN REF. 6).
 K -> R (IN REF. 12).
 SE -> ES (IN REF. 6).
 Query Match 79.0%; Score 2450.5; DB 1; Length 607;
 Best Local Similarity 75.8%; Pred. No. 4.1e-152;
 Matches 442; Conservative 71; Mismatches 69; Indels 1; Gaps 1;
 QY 1 DAHKEVAHFRFDLGEENFKALVIAFAOYLQCCPFEDHVKLVNVEVTEFAKTCVADESAE 60
 Db 25 DTHKSEIAHFRFDLGEHFKGLVLIASFQYLQCCPFEDHVKLVNVEVTEFAKTCVADESHA 84
 QY 61 NCDKSLHTLFGDKLCTVATLRETYGEMADCCAKPEPNECEFLQHKDDNPILPRLVREV 120
 Db 85 GCKSLHTLFGDELCCKVASLRETYGMDCCCKPEPNECEFLSHKDDSPDFDKL-KPDP 143
 QY 121 DVNCTAFNDHNETFLKLYEIAARRHPYFAPELLFFAKRYKAATFECQQAADKAACILP 180
 Db 144 NTLCDFADEKFKWGLYKLYEIAARRHPYFAPELLFYANKYGVFQCCQAEKGCALP 203
 QY 181 KLDELDEKSKASSAKQRLKCAKQKGEAFKNAVARLSQRPKFAEFVSKLWTLTK 240
 Db 181 KLDELDEKSKASSAKQRLKCAKQKGEAFKNAVARLSQRPKFAEFVSKLWTLTK 240

```

RESULT 7
ALBU_RAB1T
ID ALBU_RAB1T STANDARD; PRT; 508 AA.
AC P49065;
DC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Serum albumin precursor.
GN ALB.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Sheffield W.P., Syed S., Schuyler P.D.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC !- FUNCTION: Serum albumin, the main protein of plasma, has a good
CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
CC hormones, bilirubin and drugs. Its main function is the regulation
CC of the colloidal osmotic pressure of blood.
CC !- SUBCELLULAR LOCATION: Secreted.
CC !- TISSUE SPECIFICITY: Plasma.
CC !- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC !- SIMILARITY: Contains 3 albumin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U18344; AAB56347.1; -.
CC HSP; P02768; IEF7B.
CC
CC InterPro; IPR000264; Serum_albumin.
CC Pfam; PF00273; transport_prot; 3.
CC PRINTS; PR00802; SERUMALBUMIN.
CC ProDom; PD002486; Serum_albumin; 1.
CC SMART; SM00103; ALBUMIN; 3.
CC PROSITE; PS00212; ALBUMIN; 3.
CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.
KW SIGNAL 1 18
FT PROPEP 19 24 BY SIMILARITY.
FT PROPEP 19 24 BY SIMILARITY.

```


RL Nucleic Acids Res. 16:9045-9045(1988).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Plasma.

CC -!- TISSUE SPECIFICITY: Secreted.

CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; X12422; CAA30970.1; -

CC EMBL; M36787; AAA30988.1; -

CC PIR; S01382; ABPGS.

CC HSP; P02768; 1B7H.

CC InterPro: IPR000264; Serum_albumin.

CC Pfam: PF00273; transport_prot; 3.

CC PRINTS; P00802; SERUMALBUMIN.

CC PRODOM; PD002486; Serum_albumin; 1.

CC SMART; SM00103; ALBUMIN; 3.

CC PROSITE; PS00212; ALBUMIN; 3.

CC Metal-binding; Lipid-binding; Repeat; Signal; Copper.

FT NON_TER 1 1

FT SIGNAL 1 16 BY SIMILARITY.

FT PROPEP 17 22 BY SIMILARITY.

FT CHAIN 23 605 SERUM ALBUMIN.

FT DOMAIN 23 202 ALBUMIN 1.

FT DOMAIN 209 394 ALBUMIN 2.

FT DOMAIN 401 592 ALBUMIN 3.

FT METAL 31 31 COPPER (BY SIMILARITY).

FT DISULFID 75 84 BY SIMILARITY.

FT DISULFID 97 113 BY SIMILARITY.

FT DISULFID 112 123 BY SIMILARITY.

FT DISULFID 145 190 BY SIMILARITY.

FT DISULFID 189 198 BY SIMILARITY.

FT DISULFID 221 267 BY SIMILARITY.

FT DISULFID 266 274 BY SIMILARITY.

FT DISULFID 286 300 BY SIMILARITY.

FT DISULFID 299 310 BY SIMILARITY.

FT DISULFID 337 382 BY SIMILARITY.

FT DISULFID 381 390 BY SIMILARITY.

FT DISULFID 413 459 BY SIMILARITY.

FT DISULFID 458 469 BY SIMILARITY.

FT DISULFID 482 498 BY SIMILARITY.

FT DISULFID 497 508 BY SIMILARITY.

FT DISULFID 535 580 BY SIMILARITY.

FT DISULFID 579 588 BY SIMILARITY.

FT CONFLICT 562 562 E -> D (IN REF. 1; AAA30988).

SQ SEQUENCE 605 AA; 69410 MW; 3E56B0DD1A1FF CRC64;

Query Match 77.7%; Score 2411.5; DB 1; Length 605;

Best Local Similarity 76.0%; Pred. No. 1.4e-149;

Matches 439; Conservative 67; Mismatches 70; Indels 1; Gaps 1;

QY 1 DAHSEVAHRFDKIDENFALVLIAPAOYLQCCPFEDHVKLVNVEYFAKTCVADSAE 60

DB 23 DTKKSIARFNDLSEQYFGVLVLIASFQHLQCCPFYEHVKLVNVEYFAKTCVADSAE 82

QY 61 NCKDSHTLFGDKLKTATLVREYGMADCCAKOPEPNECFLOHKDDPNLPRLVPEV 120

DB 83 NCKDSHTLFGDKLCAIPSEYREHYGLADCCKEPEPNECFLOHNDPNLPKL-KDP 141

QY 121 DVMTAFHNDNEETFKLYLYEIARRHPYFAPLIFTAKRYKAAFTCCQAADKAACLLP 180

DB 142 VALCADQFDEQKFWGKLYEIARRHPYFAPLIPYLYIYKDFVSECCQAADKAACLLP 201

QY 181 KLDELDEGKASSAKORLKCASIQKQGERAFKAWAVARLSQRPKAEFAEYSLVTLTK 240

DB 202 KIEHLREKVTSAKORLKCASIQKQGERAFKAWAVARLSQRPKAEFAEYSLVTLTK 261

QY 241 VTECHGDLECADRADLAKYICENODSISSSKLEKCEKPLLEKSKCIAEVENDEMPA 300

DB 262 VHKECCHGDLECADRADLAKYICENODTISTIKLEKCEKPLLEKSKCIAEAKRDELPA 321

QY 301 DLPSLAADFVESKDYCKNAEAKDYFLGMYFYVARHPDYSVWLLRLAKTYFTLEKC 360

DB 322 DLPLEHDFVDEKVKCKNKEADYFLGTFYFSRHPDYSVWLLRLAKTYFTLEKC 381

QY 361 CAAADPHCYAKVDFEKFPLVEEPQNLKONCELFQOLGEYFQNALVRYTKKVPQVST 420

DB 382 CAKEDPPACYATVDFKFPQVLDPEKNLKONCELFQOLGEYFQNALVRYTKKVPQVST 441

QY 421 PTLVEVSRNLKGVSKCKCKHFAKMPCAEDYLSVNLQCLVLEHKTVPVSDRVTKCCTES 480

DB 442 PTLVEVARKLGLVSGRCCKPEERLSCAEDYLSVNLQCLVLEHKTVPVSEKVTCKCTES 501

QY 481 LVNRRPCFSALEVDYTVPKFNAETFTFHADICTLSEKQIKKOTALVELVKKPKAT 540

DB 502 LVNRRPCFSALTPTDPTVKPEFVEGTFTFHADLCTLPEDKQIKKOTALVELLAKHPAT 561

QY 541 KEOLKAVMDFAAEVKECKCKADDKTCFAEKGKLV 576

DB 562 EEQLRTVLGNFAAFVQKCAAPDHEACFAGVGRFV 597

RESULT 11

ALBU_MERUN STANDARD; PRT; 609 AA.

AC Q35090;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Serum albumin precursor.

GN ALB.

OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;

OC Meriones.

OX NCBI_TaxID=10047;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MGS IDR; TISSUE=Liver;

RX MEDLINE=98116663; PubMed=9455485;

RA Yoshida K., Seto-Ohshima A., Sinohara H.;

RT "Sequencing of cDNA encoding serum albumin and its extrahepatic

RL synthesis in the Mongolian gerbil, Meriones unguiculatus.";

RL DNA Res. 4:351-354(1997).

CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good

CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,

CC hormones, bilirubin and drugs. Its main function is the regulation

CC of the colloidal osmotic pressure of blood.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Plasma.

CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.

CC -!- SIMILARITY: Contains 3 albumin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; AB006197; BAA21765.1; -

DR PIR; JC5838; JC5838.

DR HSP; P02768; 1B7B.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot; 3.

DR PRINTS: PRO0802; SERUMALBUMIN.
 DR ProDom: PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18 BY SIMILARITY.
 FT PROPEP 19 24 BY SIMILARITY.
 FT CHAIN 25 609 SERUM ALBUMIN.
 FT DOMAIN 25 206 ALBUMIN 1.
 FT DOMAIN 213 398 ALBUMIN 2.
 FT DOMAIN 405 596 ALBUMIN 3.
 FT METAL 28 28 COPPER.
 FT DISULFID 78 87 BY SIMILARITY.
 FT DISULFID 100 116 BY SIMILARITY.
 FT DISULFID 115 126 BY SIMILARITY.
 FT DISULFID 149 194 BY SIMILARITY.
 FT DISULFID 193 202 BY SIMILARITY.
 FT DISULFID 225 271 BY SIMILARITY.
 FT DISULFID 270 278 BY SIMILARITY.
 FT DISULFID 290 304 BY SIMILARITY.
 FT DISULFID 303 314 BY SIMILARITY.
 FT DISULFID 341 386 BY SIMILARITY.
 FT DISULFID 385 394 BY SIMILARITY.
 FT DISULFID 417 463 BY SIMILARITY.
 FT DISULFID 462 473 BY SIMILARITY.
 FT DISULFID 486 502 BY SIMILARITY.
 FT DISULFID 501 512 BY SIMILARITY.
 FT DISULFID 539 584 BY SIMILARITY.
 FT DISULFID 583 592 BY SIMILARITY.
 SQ SEQUENCE 609 AA; 69940 MW; 9CA5F97F67FE1A48 CRC64;
 Query Match
 Best Local Similarity 76.9%; Score 2387; DB 1; Length 609;
 Matches 430; Conservative 65; Mismatches 87; Indels 0; Gaps 0;
 QY 2 AKSVVAHRKDGLENFALVLIAPAQYLOQCFPHVLYNEVEFAKTCVADSAEN 61
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 27 AHKSIARHYKDGKFGYGLVLYTFYSYLOKQCYEEHVKLVREYDFASNCADSAEN 86
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 62 CDKSLHTLFGKLTAVATREYVGENMADCCAKQPERNECFLOKDDNPNFLVPRPEVD 121
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 87 CDKSLHTLFGKLSLPNGFKYAEAMADCCAKQPERNECFLOKDDNPNFLVPRPEVD 146
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 122 VMCFAFHNETFLKYLVEIARRHPYFAPPELLFFAKRYKAFTTCCOAAKKAACLLPK 181
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 147 AMCTAFOENAEAFMGHYLHEVARHPHYFGPELLYLDKTYAVLTCCCAADDKGACLTPK 206
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 182 LDELDEKASAKQRLKASLQKGFGRFAKAVARLSORPEKAEFAVSKLYTDLTKV 241
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 207 LDALKEKALSVNRQRLKSSMKFGGERAFKAVARMSQTFPNDAFETIKLATDTKV 266
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 242 HTECHGDLLECDRADLAKYICENQDSISSKLKCECERPLEKSHGICAEVENDMPAD 301
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 267 TQECCHGDLLECDRAELAKYICENQDSISSKLKCECERPLEKSHGICAEVENDMPAD 326
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 302 LPSLAADVSKDYCKNYAEAKDVLGFLMFLYVARRPDSVLLRLAKYETYLEKCC 361
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 327 LPALTADVEDKDYCKNYAEAKDVLGFLMFLYVARRPDSVLLRLAKYETYLEKCC 386
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 362 AAADPEHYAKVDFEKLVEPEPNLITKONCELPQGLGEYKFNALLVRYTKYPOYSTP 421
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 387 AEADPHACYGVDFEKLVEPEPNLITKONCELPQGLGEYKFNALLVRYTKYPOYSTP 446
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 422 TLVFSYRNGLKVGSKCKHPEAKRMPCAEDYLSVLLNQLCVLHETKPPVSDRVTKCTESL 481
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 447 TLVEAARSLGRVGHCCALPEKRLPCVEDYLSALLNRVCLLHEKTPVSEQVTKCCSGSL 506
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 482 VNRRCFSALYVDYTYVKEFNAETFTTHADICTLSKERQIKKQATLVELVKKHPKATK 541
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 507 VERRCFCSALYVDYTYVKEFNAETFTTHADICTLSKERQIKKQATLVELVKKHPKATK 566
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||
 QY 542 EOLKAVMDFAAFVEKCKADDKETCFAEKGKLVAAASQAL 583
 DB |||||:||||: || ||:||||: ||:||||: ||||| ||||| |||||

Db 567 EQLKKYMGDFAEFLKCKQKQEDKACFTSTEGPKLVAESQKAL 608

RESULT 12

ALBU_MOUSE
 ID ALBU_MOUSE STANDARD; PRT; 608 AA.
 AC P07724; Q61802;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB OR ALBI OR ALB-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA van Reeth T., Gabant P., Dreze P., Szpirer J., Szpirer C.;
 RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Pynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RL "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 409:685-690(2001).
 RP SEQUENCE OF 99-516 FROM N.A.
 RX MEDLINE=89216123; PubMed=2452956;
 RA Minghetti P.P., Law S.W., Dugaiczky A.;
 RL "The rate of molecular evolution of alpha-fetoprotein approaches that
 of pseudogenes.";
 RN Mol. Biol. Evol. 2:347-358(1985).
 RP SEQUENCE OF 477-551 FROM N.A.
 RX STRAIN=BA1B/Cg;
 RX MEDLINE=90269606; PubMed=1971802;
 RA Boccaccio C., Deschatrette J., Meunier-Rotival M.;
 RL "Empty and occupied insertion site of the truncated LINE-1 repeat
 located in the mouse serum albumin-encoding gene.";
 RN Gene 88:181-186(1990).
 RP SEQUENCE OF 25-44.
 RC TISSUE=Liver;
 RX Glometti C.S., Taylor J., Tollaksen S.L.;
 RA "Mouse liver protein database: a catalog of proteins detected by two-
 dimensional gel electrophoresis.";
 RL Electrophoresis 13:970-991(1992).
 CC -!- FUNCTION: Serum albumin, the main protein of plasma, has a good
 of the colloidal osmotic pressure of blood.
 CC -!- SUBCELLULAR LOCATION: Secreted.

CC -- TISSUE SPECIFICITY: Plasma.
 CC -- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ011413; CAA09617.1; -
 DR EMBL; M16111; AAA37150.1; -
 DR EMBL; X13060; CAA31458.1; -
 DR EMBL; AK010025; BAB26650.1; -
 DR PIR; A05139; A05139.
 DR HSSP; P02768; 1E7B.
 DR SWISS-2DPAGE; P07724; MOUSE.
 DR MGD; MGI:87991; Alb1.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERDALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 KW Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 24
 FT CHAIN 25 608
 FT DOMAIN 25 205
 FT DOMAIN 212 397
 FT DOMAIN 404 595
 FT METAL 27 27
 FT DISULFID 77 86
 FT DISULFID 99 115
 FT DISULFID 114 125
 FT DISULFID 148 193
 FT DISULFID 192 201
 FT DISULFID 224 270
 FT DISULFID 269 277
 FT DISULFID 289 303
 FT DISULFID 302 313
 FT DISULFID 340 395
 FT DISULFID 384 393
 FT DISULFID 416 462
 FT DISULFID 461 472
 FT DISULFID 485 501
 FT DISULFID 500 511
 FT DISULFID 538 583
 FT DISULFID 582 591
 FT CONFLICT 27 27
 FT CONFLICT 33 33
 FT CONFLICT 41 41
 SQ SEQUENCE 608 AA; 68692 MW; 292F7C7EED3A61B4 CRC64;
 Query Match 76.6%; Score 2378; DB 1; Length 608;
 Best Local Similarity 72.4%; Pred. No. 2.1e-147;
 Matches 422; Conservative 80; Mismatches 81; Indels 0; Gaps 0;
 QY 1 DAHKSVAHRFDLGEENFKALVLIAPAOYLQCPDFDHVKLVNVEVFPAKTCVADSAE 60
 DB 25 EAHKSEIAHYNDLGEHFGLVLIAPSOYLQCSYDEHAKVIOEVDFPAKTCVADSAE 84
 QY 61 NCKSLHTLFGDKLTVALRETYGENADCCAKQEPERNECFLOHKDDNPLRLPRPV 120
 DB 85 NCKSLHTLFGDKLCAIPNLRENYGELADCCQEPERNECFLOHKDDNPLRPRPEA 144
 QY 121 DVMTAFPHNEETFLKLYLRIARRHPFYFAPPELLFAKRYKAFTSCQAADKACLIP 180
 DB 145 EAMCTGFKNPTFMGHVLRVARRHPFYFAPPELLFYAFQYNEILTQCQAEADKESCLTP 204
 QY 161 KLDELDEGKASSAKORLKCASLQKFGERAFAKAWARLSQRFPAEFAVSKLVITLTK 240

DB 205 KLDGVKRALVSVKQRMKSSMKQFGERAFKAWARLSQTFPNADFAEITKLAIDLTK 264
 QY 241 VITECHGDLEACADDLAKYICENODISSKLCCKECPILLESKHGCIARVENDEMPA 300
 DB 265 VKECHGDLEACADDLAKYICENQATISSKLOTCCKPILKKAHLSVEHDTMPA 324
 QY 301 DLPSLAADFVSKDCKVKNYAEAKDFVLMGYEYARHPDYSVLLRLAKYETTLK 360
 DB 325 DLPALAADFVEDQVCKNYAEAKDFVLMGYEYARHPDYSVLLRLAKYETTLK 384
 QY 361 CAAADPHCYAKVPEKPIVEEPQNLKONCELFQELGEYKTONALLYRYKKYPOYST 420
 DB 385 CAEAPPACYGTVAEFPQVLEEFKPNLNCDLTKELGEYKTONALLYRYKKYPOYST 444
 QY 421 PTLVEVSRNLGKVGSKCKPEAKRMPCAEDYLSVVLNQLCVLHEKTPVSDRVTKCCTES 480
 DB 445 PTLVEAARNLGRVGTCKCTLPEORLPCEVDYLSAILNRVCLLHEKTPVSEHVKCCSGS 504
 QY 481 LYNRPCTSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKQATLVLYVKKHKPAT 540
 DB 505 LVERRPCTSALEVDETYVPKEFNAETFFHADICTLSEKEROIKKQATLVLYVKKHKPAT 564
 QY 541 KEOLKAVMDRAAFVKECKCKADDOCKTCFAEKGKLVLAASQAAL 583
 DB 565 AEQLKAVMDRAAFVKECKCKADDOCKTCFAEKGKLVLAASQAAL 607
 RESULT 13
 ID ALBU_CHICK STANDARD; PRT; 615 AA.
 AC P19121;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Serum albumin precursor.
 GN ALB.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Cassidy A.I., Salkild C.K., Baverstock P., Wallace J.C.;
 RN Submitted (JUL-1991) to the EMBL/GenBank/DDJB databases.
 [2]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=83161037; PubMed=6187737;
 RA Hache R.J.G., Wiskocil R., Vasa M., Roy R.N., Lau P.C.K., Deeley R.G.;
 RT "The 5' noncoding and flanking regions of the avian very low density
 RT apolipoprotein II and serum albumin genes. Homologies with the egg
 RT white protein genes".
 RL J. Biol. Chem. 258:4556-4564(1983).
 RN [3]
 RP SEQUENCE OF 19-30.
 RX MEDLINE=78019943; PubMed=911327;
 RA Rosen A.M., Geller D.M.;
 RT "Chicken microsomal albumin: amino terminal sequence of chicken
 RT proalbumin".
 RL Biochem. Biophys. Res. Commun. 78:1060-1066(1977).
 CC -- FUNCTION: Serum albumin, the main protein of plasma, has a good
 CC binding capacity for water, Ca(2+), Na(+), K(+), fatty acids,
 CC hormones, bilirubin and drugs. Its main function is the regulation
 CC of the colloidal osmotic pressure of blood.
 CC -- SUBCELLULAR LOCATION: Secreted.
 CC -- TISSUE SPECIFICITY: Plasma.
 CC -- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Y60688; CAA43098.1; --
 DR EMBL; V00381; CAA23680.1; --
 DR PIR; S15571; ABCHS.
 DR HSP; P02768; 1E7B.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 3.
 DR Metal-binding; Lipid-binding; Repeat; Signal; Copper.
 FT SIGNAL 1 18
 FT PROPEP 19 23
 FT CHAIN 24 615 SERUM ALBUMIN.
 FT DOMAIN 24 209 ALBUMIN 1.
 FT DOMAIN 216 401 ALBUMIN 2.
 FT DOMAIN 408 599 ALBUMIN 3.
 FT METAL 30 30 COPPER (BY SIMILARITY).
 FT DISULFID 80 89 BY SIMILARITY.
 FT DISULFID 102 118 BY SIMILARITY.
 FT DISULFID 117 128 BY SIMILARITY.
 FT DISULFID 152 197 BY SIMILARITY.
 FT DISULFID 196 205 BY SIMILARITY.
 FT DISULFID 228 274 BY SIMILARITY.
 FT DISULFID 273 281 BY SIMILARITY.
 FT DISULFID 293 307 BY SIMILARITY.
 FT DISULFID 306 317 BY SIMILARITY.
 FT DISULFID 344 389 BY SIMILARITY.
 FT DISULFID 388 397 BY SIMILARITY.
 FT DISULFID 420 466 BY SIMILARITY.
 FT DISULFID 465 476 BY SIMILARITY.
 FT DISULFID 489 505 BY SIMILARITY.
 FT DISULFID 504 515 BY SIMILARITY.
 FT DISULFID 542 587 BY SIMILARITY.
 FT DISULFID 586 595 BY SIMILARITY.
 FT CARBOHYD 500 500 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CONFLICT 24 24 F -> M (IN REF. 3).
 SQ SEQUENCE 615 AA; 69918 MW; E59E4BCEC06C6 CRC64;
 Query Match 50.2%; Score 1557.5; DB 1; Length 615;
 Best Local Similarity 46.7%; Pred. No. 4.3e-94;
 Matches 273; Conservative 118; Mismatches 192; Indels 1; Gaps 1;
 QY 3 HKSEVAHFKDLEENFKALVLIATFALQVLOQCPEDHVKLVNEVEFAKTCVADSEANBC 62
 DB 30 HKSEIAHRYNDLKEETFKAVAMITFAQLQRCYSGLSKLVQDVLDLQCKVANEADPEC 89
 QY 63 DKSLLTFLGDKLCTVATLRETYGEMADCAKOEPEKNECFLOKHQNDPNLPR-LVPRVD 121
 DB 90 SKFLPSIIILDCQVEKLRDSYGAMADCCSKADPERNECTLSFKVSQPFVQYORASD 149
 QY 122 VMCTAFHNEFTFKLYLFIARHPHYFVAPELLFFAKRYKAFTCCQAADKAACLIPK 181
 DB 150 VICQEQDNRVSVFLGHETYSVARRHPFLYAPALLSFAYDFEHALQSCCKESDVGACLDTK 209
 QY 182 LDELDEKASAKORLKCSAQKCFERAKAWARLSORFPKAEFAVSKLVLDLTKY 241
 DB 210 EIVMREKAGSVSQYFCGLKQFGDRVQARQIYLSQKPKAPFSEVSVFVHDSIGV 269
 QY 242 HFECHGDLLECCADRALAKYICENQDSISKLECEKPLLEKSHCHIAVEVNDENP 301
 DB 270 HKECCGDMVCMDDMAMNSNLCSSQDVFSGKINDCCERPIVERSQCIWEAFDEKPAD 329
 QY 302 LPSLAADFVESKQVKNYAKADKDFVGLMFLYETARRHPDYVSWLLLRKATYETLEKCC 361
 DB 330 LPSLVKYEIEDKEVCKSFEGADFAFVAEVIYSRRHPEFSIQLIMRIAKGESLLEKCC 389

QY 362 AADPHECYAKVDFEKFPLVEEPQNLKQNCLEFQOLGEYFQNALLVRYTKKVPQVSTP 421
 DB 390 KTDNPAECYANAOEQUNQHIKETQBVVYKTNCDLHDHGEADFLKSILIRTKKMPQVPTD 449
 QY 422 TLVSVSNLKGVSCKCKHPEAKRMPCAEDYLSVLNQLCVLHETKTPVSDRVTKCCTESTL 481
 DB 450 LLLTGGKWTIGTKCCQGLGEDRMACEGYLSIVIHDTCKRQETTPINDVNSQCCSLY 509
 QY 482 VNRPCPSALEVDYVYKFNATFTTHADICTLSEKREQIKKQTAIVELVHKHKAUK 541
 DB 510 ANRPPCTAMGVDYTKYVPPFPNDFMSFDEKLSAFAPAEEREVGQMKLLINLIRKQPMTE 569
 QY 542 EQLKAVMDDFAAVFEKCKKADDKETCFAEKGGKLVAAASOALGL 585
 DB 570 EQIKTIADGFTAMVDKCKQSDINFCFGEAGANLIVOSRNLGI 613
 RESULT 14
 PETA_PANTR
 ID PETA_PANTR STANDARD; PRG; 609 AA.
 AC Q28789.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-fetoprotein).
 DE DE
 GN AFP.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
 OC NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96032345; PubMed=7557431;
 RA Nishio H., Gibbs P.E., Minghetti P.P., Zielinski R., Dugaiczak A.;
 RT "The chimpanzee alpha-fetoprotein-encoding gene shows structural
 RT similarity to that of gorilla but distinct differences from that of
 RT human";
 RL Gene 162:213-220(1995).
 CC -!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND
 CC BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE
 CC (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
 CC -!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION
 CC TO THE MONOMERIC FORM.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Plasma. Synthesized by the fetal liver and
 CC yolk sac.
 CC -!- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -!- SIMILARITY: Contains 3 albumin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

EMBL; U21916; AAA91641.1; -

DR PIR; JC4258; JC4258.

DR HSP; P02768; 1E7B.

DR InterPro; IPR000264; Serum_albumin.

DR Pfam; PF00273; transport_prot; 3.

DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum_albumin; 1.

DR SMART; SM00103; ALBUMIN; 3.

DR PROSITE; PS00212; ALBUMIN; 2.

KW Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
 KW Signal.

FT SIGNAL 1 18 BY SIMILARITY.

FT CHAIN 19 609 ALPHA-FETOPROTEIN.

FT DOMAIN 20 205 ALBUMIN 1.

FT DOMAIN 212 397 ALBUMIN 2.

```

FT DOMAIN 404 595 ALBUMIN 3.
FT METAL 22 22 COPPER AND NICKEL (BY SIMILARITY).
FT DISULFID 99 114 BY SIMILARITY.
FT DISULFID 113 124 BY SIMILARITY.
FT DISULFID 148 193 BY SIMILARITY.
FT DISULFID 192 201 BY SIMILARITY.
FT DISULFID 224 270 BY SIMILARITY.
FT DISULFID 269 277 BY SIMILARITY.
FT DISULFID 289 303 BY SIMILARITY.
FT DISULFID 302 313 BY SIMILARITY.
FT DISULFID 384 393 BY SIMILARITY.
FT DISULFID 416 462 BY SIMILARITY.
FT DISULFID 461 472 BY SIMILARITY.
FT DISULFID 485 501 BY SIMILARITY.
FT DISULFID 500 511 BY SIMILARITY.
FT DISULFID 538 583 BY SIMILARITY.
FT DISULFID 582 591 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 609 AA; 68741 MW; C032987CADOE672B CRC64;

Query Match 40.4%; Score 1253.5; DB 1; Length 609;
Best Local Similarity 40.1%; Pred. No. 2.4e-74;
Matches 236; Conservative 115; Mismatches 231; Indels 7; Gaps 3;

QY 3 HKSE-----VAHFEDLGEENFKALVLIAPAQVLOQCPEDEAVKLYNEVTEFAKTCVAD 57
DB 22 HRNEYTASLDSTQCTACINLTDLATIFFAQVQVATYKEVSKMVKDALTAEKPTGDE 81
QY 58 SAENCKSLTFEGDKLCVAVTLRETYGEMADCCAQEPNECFLOKHQDNP-NLPRLV 116
DB 82 QSAGLENQLPALFELCEKEILEKYGH-SDCCSQSEBGRHNCFLAHKPTPASIFPQ 140
QY 117 RPEVDVMTAFHNETFLKYLVEIARRHPYIYAPPELLFFAKRYKAFAFECQADKAA 176
DB 141 VPPEVTCEAYEDREFTFNKFTYIARRHPFTYAPTILLWAARYDKLIPSCCKAENAYE 200
QY 177 CLLPKDELDEKASKAKORLKCASQKQFGERAFKAWAVARLSORFFKAFBVSRLVT 236
DB 201 CEQPKATVYKELRESLILNHACAVMKNFTFTFOATVYKLSQKFKVNFIEQLVL 260
QY 237 DLTVHTFCHGDLBCADRADLAKYICENQDISKILKECCKEPLEKSHCHTAEVND 296
DB 261 DVAVHHEHCCRGVDLCLQDGEKIMSYTCQQDFTLKNKITECKLTTLTGQCIIHAEND 320
QY 297 EMPADLSLAADVESKDVCKNYAEKDVFLGMELEYARRHPDYSVLLILRLAKTYET 356
DB 321 EKPGLSPNRLTGDRDNQFSSEKRNIFLASFVHEISRRPQLAVSVILRVKANGQEL 380
QY 357 LEKCAAADPHECYAKVFDEPKPLVEEPQNLIKQNCSELFQELGEYKFFQALLVRYTKVP 416
DB 381 LEKCFQTEPLEQDQGEELQYIQESQALAKRSCLGFKLGEYLLQNAFLVAYTKAP 440
QY 417 QYSTPTLVEVSNLGVSKCKHPEAKMECAEDYLSVVLNQLCVLHEKTPSDRYTKC 476
DB 441 QUTSELMAITRWKATAATACQLSEDKLLACGGGAADIILGHICIRHETPNVPGVQC 500
QY 477 CTESIVNRPCFSALVEVDYTPYKFEFNAETFTFHADICTLSEKQIKKQATALVELYKHK 536
DB 501 CTSSYANRRPCFSLVVDYTPYPAFSDDKFIHFKDLCOAQGVALTMTKQEFLLINVKQK 560
QY 537 PKATKEQLKAVMDFAAFVEKCKADDKETCFABEGKLVVAASQAALG 585
DB 561 PQITBEQLEAVIADFSLGLEKCCGQEQEVCFABEGOKLISKTRALGV 609

```

RESULT 15

PETA_HUMAN

ID PETA_HUMAN STANDARD; PRT; 609 AA.

AC P02771;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
DE fetoprotein).
DE GN AFF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=83273664; PubMed=6192439;
RX Morinaga T., Sakai M., Wegmann T.G., Tamaoki T.;
RT "Primary structures of human alpha-fetoprotein and its mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:4604-4608(1983).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=87185438; PubMed=2436661;
RX Gibbs P.E.M., Zielinski R., Boyd C., Dugaiczky A.;
RT "Structure, polymorphism, and novel repeated DNA elements revealed by
a complete sequence of the human alpha-fetoprotein gene.";
RL Biochemistry 26:1332-1343(1987).
RN [3]
RN SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=93278385; PubMed=7684942;
RX McVey J.H., Michaelides K., Hansen L.P., Ferguson-Smith M.,
RA Tilghman S., Krumlauf R., Tuddenham E.G.D.;
RT "A G->A substitution in an HNF I binding site in the human alpha-
fetoprotein gene is associated with hereditary persistence of alpha-
fetoprotein (HPAFP).";
RL Hum. Mol. Genet. 2:379-379(1993).
RN [4]
RN SEQUENCE OF 429-556 FROM N.A.
RX MEDLINE=91242409; PubMed=6187626;
RX Beattie W.G., Dugaiczky A.;
RT "Structure and evolution of human alpha-fetoprotein deduced from
partial sequence of cloned cDNA.";
RL Gene 20:415-422(1982).
RN [5]
RN PARTIAL SEQUENCE OF 19-609.
RX MEDLINE=91242409; PubMed=1709810;
RX Pucci P., Scigliano R., Malorni A., Marino G., Tecce M.F.,
RA Ceccarini C., Ferrana B.;
RT "Human alpha-fetoprotein primary structure: a mass spectrometric
study.";
RL Biochemistry 30:5061-5066(1991).
RN [6]
RN PRELIMINARY SEQUENCE OF 19-35.
RX MEDLINE=77242506; PubMed=70228;
RX Tachin S., Hsu R., Heinrichson R.L., Miller J.B.;
RT "Studies on human alpha-fetoprotein. Isolation and characterization
of monomeric and polymeric forms and amino-terminal sequence
analysis.";
RL Biochim. Biophys. Acta 493:418-428(1977).
RN [7]
RN PRELIMINARY SEQUENCE OF 19-38.
RX MEDLINE=78001760; PubMed=71198;
RX Aoyagi Y., Ikenaka T., Ichida F.;
RT "Comparative chemical structures of human alpha-fetoproteins from
fetal serum and from ascites fluid of a patient with hepatoma.";
Cancer Res. 37:3663-3667(1977).
RN [8]
RN PRELIMINARY SEQUENCE OF 19-39.
RX MEDLINE=75018719; PubMed=4138095;
RX Ruoslahti E., Pihko H., Vaheri A., Seppala M., Virolainen M.,
RA Kontinen A.;
RT "Alpha fetoprotein: structure and expression in man and inbred mouse
strains under normal conditions and liver injury.";
RL Johns Hopkins Med. J. Suppl. 3:249-255(1974).
RN [9]
RN GENE STRUCTURE.
RX MEDLINE=85182629; PubMed=2580830;
RX Sakai M., Morinaga T., Urano Y., Watanabe K., Wegmann T.G.,
RA Tamaoki T.;
RT "The human alpha-fetoprotein gene. Sequence organization and the 5'

```

RT	flanking region. ";
RN	J. Biol. Chem. 260:5055-5060(1985).
RA	[10]
RX	METAL-BINDING.
RP	MEDLINE=79001617; PubMed=80265;
RA	Aoyagi Y., Ikenaka T., Ichida F.;
RT	"Copper (II)-binding ability of human alpha-fetoprotein.";
RT	Cancer Res. 38:3483-3486(1978).
RN	[11]
RN	BILIRUBIN-BINDING.
RP	MEDLINE=80001710; PubMed=89900;
RA	Aoyagi Y., Ikenaka T., Ichida F.;
RT	"Alpha-fetoprotein as a carrier protein in plasma and its bilirubin-binding ability. ";
RT	Cancer Res. 39:3571-3574(1979).
RN	[12]
RP	SULFATION.
RX	MEDLINE=86042625; PubMed=2414772;
RA	Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
RT	"Tyrosine sulfation of proteins from the human hepatoma cell line HepG2. ";
RT	Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
CC	-!- FUNCTION: BINDS COPPER, NICKEL, AND FATTY ACIDS AS WELL AS, AND BILIRUBIN LESS WELL THAN, SERUM ALBUMIN. ONLY A SMALL PERCENTAGE (LESS THAN 2%) OF THE HUMAN AFP SHOWS ESTROGEN-BINDING PROPERTIES.
CC	-!- SUBUNIT: DIMERIC AND TRIMERIC FORMS HAVE BEEN FOUND IN ADDITION TO THE MONOMERIC FORM.
CC	-!- SUBCELLULAR LOCATION: Secreted.
CC	-!- TISSUE SPECIFICITY: PLASMA. SYNTHESIZED BY THE FETAL LIVER AND YOLK SAC.
CC	-!- DEVELOPMENTAL STAGE: OCCURS IN THE PLASMA OF FETUSES MORE THAN 4 WEEKS OLD, REACHES THE HIGHEST LEVELS DURING THE 12TH-16TH WEEK OF GESTATION, AND DROPS TO TRACE AMOUNTS AFTER BIRTH. THE SERUM LEVEL IN ADULTS IS USUALLY LESS THAN 40 NG/ML. AFP OCCURS ALSO AT HIGH LEVELS IN THE PLASMA AND ASCITIC FLUID OF ADULTS WITH HEPATOMA.
CC	-!- PTM: INDEPENDENT STUDIES SUGGEST HETEROGENEITY OF THE AMINO-TERMINAL SEQUENCE OF THE NATIVE PROTEIN AND OF THE CLEAVAGE SITE OF THE SIGNAL SEQUENCE.
CC	-!- PTM: SULFATED.
CC	-!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC	-!- SIMILARITY: Contains 3 albumin domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M10949; AAA51674.1; -
DR	EMBL; M10950; AAA51675.1; -
DR	EMBL; V01514; CAA24758.1; -
DR	EMBL; M16110; AAB58754.1; -
DR	EMBL; Z19532; CAA79592.1; -
DR	PIR; A26624; FPHU.
DR	HSSP; P02768; IE7B.
DR	GlycoSuiteDB; P02771; -
DR	Siena-2DPAGE; P02771; -
DR	Genew; HGNC:317; AFP.
DR	MM; 104150; -
DR	InterPro; IPR000264; Serum_albumin.
DR	Pfam; PF00273; transport_prot; 3.
DR	PRINTS; PR00802; SERUMALBUMIN.
DR	ProDom; PD002486; Serum_albumin; 1.
DR	SMART; SM00103; ALBUMIN; 3.
DR	PROSITE; PS00212; ALBUMIN; 2.
KW	Glycoprotein; Sulfation; Repeat; Metal-binding; Copper; Nickel;
FT	Signal; Polymorphism.
FT	SIGNAL 1 18
FT	CHAIN 19 609
FT	DOMAIN 20 205
FT	DOMAIN 212 397
FT	FT

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 31, 2003, 16:33:10 ; Search time 106 Seconds
(without alignments)

1424.158 Million cell updates/sec

Title: US-09-833-041-18

Perfect score: 3103
Sequence: 1 DAHKSEVAHRFKDLGRENFK.....TCFAEKGKLVAAQAALGL 585

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organella.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertebrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp.bacteriap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2504	80.7	608	5 Q95VB7	Q95vb7 schistosoma
2	2374	76.5	608	11 Q8C7H3	Q8c7h3 mus musculus
3	2336	75.3	576	11 Q8C7C7	Q8c7c7 mus musculus
4	1865.5	60.1	396	4 Q8IUK7	Q8iuk7 homo sapien
5	1295.5	41.7	527	13 Q8JIA9	Q8jia9 sphendonon p
6	1242	40.0	609	6 Q8MJU5	Q8mju5 canis famil
7	1242	40.0	626	13 Q8UW05	Q8uw05 ambystoma m
8	1218.5	39.3	610	6 Q8MJ76	Q8mj76 sus scrofa
9	1087	35.0	624	13 Q8UW06	Q8uw06 ambystoma t
10	1082	34.9	605	11 Q8BK65	Q8bk65 mus musculus
11	1076	34.7	605	11 Q8BK56	Q8bk56 mus musculus
12	1045	33.7	400	13 Q8JIA7	Q8jia7 sphendonon p
13	955	30.8	603	13 Q8YGH6	Q8ygh6 rana shqipe
14	928.5	29.9	614	13 Q9L134	Q9l134 najja najja
15	888	28.6	406	13 Q8JIA8	Q8jia8 hoplodactyl
16	739	23.8	205	11 Q8CG74	Q8cg74 mus musculus

17	713	23.0	417	11 Q8R0J9	Q8r0j9 mus musculus
18	376.5	12.1	484	13 Q9W6F5	Q9w6f5 gallus gall
19	373	12.0	476	11 Q9CY31	Q9cy31 mus musculus
20	372	12.0	476	11 Q9IXG1	Q9ixg1 mus musculus
21	365	11.8	551	13 Q42279	Q42279 petromyzon
22	331	10.7	122	13 Q9WZ8	Q9wz8 larus argen
23	290	9.3	123	13 Q9WZ6	Q9wz6 poephila gu
24	264	8.5	135	11 Q63205	Q63205 rattus norv
25	188	6.1	1723	2 Q9JMK8	Q9jmx8 helicobacte
26	184	5.9	1819	16 Q9LIV0	Q9livo helicobacte
27	184	5.9	1927	16 Q25262	Q25262 helicobacte
28	162.5	5.2	44	6 Q95MC2	Q95mc2 equus cabal
29	162.5	5.2	680	5 Q9V6S8	Q9v6s8 drosophila
30	161	5.2	1079	3 Q96V11	Q96v11 pneumocysti
31	156	5.0	1026	3 Q74669	Q74669 pneumocysti
32	154.5	5.0	3843	5 Q9U5D0	Q9u5d0 drosophila
33	153.5	4.9	661	5 Q8MS79	Q8ms79 drosophila
34	153.5	4.9	3843	5 Q9VT94	Q9vu94 drosophila
35	153	4.9	62	6 Q8M111	Q8m111 sus scrofa
36	150	4.8	1065	3 Q01828	Q01828 pneumocysti
37	149	4.8	40	6 Q9TRA5	Q9tra5 oryctolagus
38	148	4.8	1028	3 Q74668	Q74668 pneumocysti
39	144.5	4.7	8749	4 Q8NF91	Q8nf91 homo sapien
40	142.5	4.6	2756	10 Q9LJ60	Q9lj60 arabidopsis
41	141.5	4.6	1560	5 Q26644	Q26644 strongyloc
42	138.5	4.5	1069	3 Q96VI2	Q96vi2 pneumocysti
43	137.5	4.4	1348	16 Q8YK55	Q8yk55 anabaena sp
44	135.5	4.4	2841	5 Q8MLU9	Q8mlu9 drosophila
45	135.5	4.4	2931	5 Q9W2C6	Q9w2c6 drosophila

ALIGNMENTS

RESULT 1

Q95VB7 PRELIMINARY; PRT; 608 AA.
AC Q95VB7;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Albumin.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_taxid=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Osman A., Asahi H., Staecker M.J., LoVerde P.T.;
RT "Albumin precursor homolog is a novel T helper cell immunogenic egg
RT component in murine infection with Schistosoma mansoni";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418550; AAL08579.1; --
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PF00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
DR SEQUENCE 608 AA; 68225 MW; E5EABB28E1C66E54 CRC64;

Query Match 80.7%; Score 2504; DB 5; Length 608;
Best Local Similarity 76.3%; Pred. No. 1.7e-188;
Matches 445; Conservative 79; Mismatches 59; Indels 0; Gaps 0;

QY 1 DAHKSEVAHRFKDLGRENFKALVIAFQAQYLOQCFPEDHVKLVNEVTEFAKTCVADESAAE 60
Db 25 DAHKSEIHRFKDLGEQHFGLVLIAPSQLQPCPYEHRVKLVNEVTEFAKTCVADESAAE 84
QY 61 NCKSLHTLFGKLCITVATLRETYGEMADCCAKQPERNECFLOHKDDNINLPIVRPV 120
Db 85 NCKSLHTLFGKLCIAITLREDSYGLADCCAKQPERNECFLOHKDDNINLPIVRPV 144


```

QY 249 DLLECAADRADLAKYICENQDSISSKLECECKPILKSHCIAVENDEMPADLPISLAAD 308
Db 241 DLLECAADRAELAKYICENQDSISSKLECECKPILKSHCIAVENDEMPADLPISLAAD 300
QY 309 FVSKDCKNYAABADKDFLGMFLYEVARRHPDYSVLLLRKAKYETTLKCCCAADPHE 368
Db 301 FVEQEVCKNYAABADKDFLGMFLYEVARRHPDYSVLLLRKAKYETTLKCCCAADPHE 360
QY 369 CYAKVDFEKPILVEEPQNLKONCELFQELGEYKFNALYVRYTKKYPQVSTPTLVEVSR 428
Db 361 CYGVLAEPQILVEEPQNLKONCELFQELGEYKFNALYVRYTKKYPQVSTPTLVEVSR 420
QY 429 NLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTPVS DRVTKCCTESLVNRPQF 488
Db 421 NLGRVTKCTLPEDQRLPCVEDYLSALNRLVCLLHEKTPVSEHVTCCSGSVLVRPQF 480
QY 489 SALEVDETYVPKEFNAETFTPHADICTLSEKQRIKKQTAALVELVKKPKATKEQLKATM 548
Db 481 SALTVDETYVPKEFNAETFTPHADICTLSEKQRIKKQTAALVELVKKPKATKEQLKATM 540
QY 549 DPAFAVEKCCADKDKTCFAEKGKLVVAASQAAL 583
Db 541 DPAFAVEKCCADKDKTCFAEKGKLVVAASQAAL 575

RESULT 4
Q81UK7 PRELIMINARY: PRT; 396 AA.
ID Q81UK7
AC Q81UK7;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DE Similar to serum albumin precursor.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC035969; AAH35969.1;
SQ SEQUENCE 396 AA; 45159 MW; 756519C096463A9B CRC64;

```

```

Query Match 60.1%; Score 1865.5; DB 4; Length 396;
Best Local Similarity 63.6%; Pred. No. 1.6e-138;
Matches 372; Conservative 0; Mismatches 0; Indels 213; Gaps 1;

QY 1 DAHSEVAHRFDLGEENFKALVLIAPAFQYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 60
Db 25 DAHSEVAHRFDLGEENFKALVLIAPAFQYLOQCPEDHVKLVNEVTEFAKTCVADSEAE 84
QY 61 NCKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 120
Db 85 NCKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 144
QY 121 DVMTAFPHNEETFLKKYLIEARRHPYFYAPPELLFAKRYKAFTCCQAADKACLIP 180
Db 145 DVMTAFPHNEETFLKKYLIEARRHPYFYAPPELLFAKRYKAFTCCQAADKACLIP 163
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSQRPKABFAEVSKLVDTLK 240
Db 164 ----- 163
QY 241 VHTECHGDLLECAADRADLAKYICENQDSISSKLECECKPILKSHCIAVENDEMPA 300
Db 164 ----- 163
QY 301 DLPSLAADFVSKDCKNYAABADKDFLGMFLYEVARRHPDYSVLLLRKAKYETTLKCC 360
Db 164 ----- 163
QY 361 DLPSLAADFVSKDCKNYAABADKDFLGMFLYEVARRHPDYSVLLLRKAKYETTLKCC 360
Db 164 ----- 163

```

```

QY 361 CAADAPHECYAKVDFEKPILVEEPQNLKONCELFQELGEYKFNALYVRYTKKYPQVST 420
Db 172 CAADAPHECYAKVDFEKPILVEEPQNLKONCELFQELGEYKFNALYVRYTKKYPQVST 231
QY 421 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTPVS DRVTKCCTES 480
Db 232 PTLVEVSRNLGKVGSKCKHPEAKRMPCAEDYLSVNLQCLVLEHTPVS DRVTKCCTES 291
QY 481 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKQRIKKQTAALVELVKKPKAT 540
Db 292 LVNRRPCFSALEVDETYVPKEFNAETFTPHADICTLSEKQRIKKQTAALVELVKKPKAT 351
QY 541 KEOLKAVMDPAFAVEKCCADKDKTCFAEKGKLVVAASQAALGL 585
Db 352 KEOLKAVMDPAFAVEKCCADKDKTCFAEKGKLVVAASQAALGL 396

RESULT 5
Q81IA9 PRELIMINARY: PRT; 527 AA.
ID Q81IA9
AC Q81IA9;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Serum albumin (Fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J.; Brennan S.O.; George P.M.; Chambers G.K.;
RT "Partial coding sequence for Sphenodon punctatus 68 kDa albumin.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF375971; AAH46104.1;
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD082486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
FT NON_TER 1
SQ SEQUENCE 527 AA; 59711 MW; C62B799E387F5929 CRC64;

```

```

Query Match 41.7%; Score 1295.5; DB 13; Length 527;
Best Local Similarity 45.9%; Pred. No. 1.4e-93;
Matches 240; Conservative 92; Mismatches 188; Indels 3; Gaps 3;

QY 62 CDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 120
Db 5 CDKSLHTLFGDKLTATLRETYGEMADCCAKQEPERNECFLOHKDDNPNLPLRVPRV 63
QY 121 DVMTAFPHNEETFLKKYLIEARRHPYFYAPPELLFAKRYKAFTCCQAADKACLIP 180
Db 64 EVSKLYQDRLTLTGNTIYEVARRHPYLVQVPPVFATASLYDEALKTCCTADKATCFHP 123
QY 181 KLDELDEGKASSAKQRLKASLQKGFGERAFKAWAVARLSQRPKABFAEVSKLVDTLK 240
Db 124 RIPPLLEYLKMNGIQENTCGILKFGERTLAKTLKLVQMSQKFPKADFAINKLVEDITH 183
QY 241 VHTECHGDLLECAADRADLAKYICENQDSISSKLECECKPILKSHCIAVENDEMPA 300
Db 184 MHTECCRGDTLLECDRADLAKYICENQDSISSKLECECKPILKSHCIAVENDEMPA 243
QY 301 DLPSLAADFVSKDCKNYAABADKDFLGMFLYEVARRHPDYSVLLLRKAKYETTLKCC 360
Db 244 DLPSLAADFVSKDCKNYAABADKDFLGMFLYEVARRHPDYSVLLLRKAKYETTLKCC 303
QY 361 CAADAPHECYAKVDFEKPILVEEPQNLKONCELFQELGEYKFNALYVRYTKKYPQVST 420
Db 304 CKTDNPECYGAAEDLKKHIAFOELVQONCDNLTLGLYFLNALLIRYTKRMPQLTS 363

```


	: : : : : :	
Db	337 DPVCKPFEEDGAPWGRFLDYAKIIPESAEINLRISAGLEKAYTCCGAAHNCEIA	396
QY	372 KVPDEFKPIVEBPONLILCONCELPEQLGEYKKFONALLRYITKKYPQVSTPLTEVSERNIG	431
Db	397 KEETLFHEITEASKLKATTCGALKELPGYHFHONIMIVRTGITLPGRSDAFLLYITKTIT	456
QY	432 KVSKCKCKPEAKRPCEADSYLVVLNOLCVLHEKPVP-SDRVTKCTCESLVNRRCFSA	490
Db	457 NIGOKCKLPDEQOMCECGGLGMVEFAQIC-QNKPTPFENEKLACHCKDSLSFTFCFAA	515
QY	491 LEVDHYVPKEFNAAETFFTHADICTLSEKPEQIKKOTALVELVKHKPKAKEQLKAVMD	550
Db	516 LTVDETVPAPVTVAESFPNFDEFCTPVSADLQAQKOTFLMHIVRTPHTIDEQVYKISEK	575
QY	551 FAAPVEKCKADDKETCFABEGKKLVA	578
Db	576 FLAMCGQCCKADORNECFTATGAKLIVA	603

RESULT	10
Q8BK65	
ID	PRELIMINARY; PRT; 605 AA.
AC	Q8BK65;
DT	01-MAR-2003 (TREMBLrel. 23, Created)
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE	Alpha-fetoprotein precursor.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCB_I_TaxID=10090;
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=C57BL/6J; TISSUE=Body;
RC	MEDLINE=223354683; PubMed=12468851;
EX	The FANTOM Consortium,
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;
RA	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs.";
RT	Nature 420:563-573(2002).
FL	EMBL; AK076053; BAC36150.1; --
DR	SEQUENCE 605 AA; 67732 MW; 048B7A48B01EA4B CRC64;
SQ	

```

Db      395  BELOKHTEESQALSKOSCALYOTLGDKYLQNLFLIGYTRKAPOLTSAEILIDTGMVSLA 454
QY      435  SKCKCHPEAKRMPCAEYLSSVLNOLVHEKTPVSDRVTKCTCTESLVNRRPCTSALEVD 494
Db      455  STCCQLSEEKWSGGGMADIIFGHLCITRNASVPVNSGISHCNCSYSNRRLCITSLFIRD 514
QY      495  EYTPYKPFNAETFTFHADICTLSEKERQIKKQATVALVEKVRKPKATKQLKAVMDFAAF 554
Db      515  EYVAPPPFSEDKETFKKDLCAQHGKALQTMKQELLINLVKQKPELTEQLAAVTDFSG 574
QY      555  VEKCKADDKTCFAPEGKIKVAAQAALGL 585
Db      575  LEKCKAQDQEVCFTEGPKLISKTRDALGV 605

RESULT 11
Q8BK56 PRELIMINARY; PRT; 605 AA.
ID Q8BK56;
AC Q8BK56;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE Alpha-fetoprotein precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
SQ EMBL; AK076197; BAC36249.1; -.
SQ SEQUENCE 605 AA; 67409 MW; CF3509A9FC146110 CRC64;

```

```

Db 455 STCCQLSEKWSGCEGMADIFIGHLCIRNEASPVNSGTSHCNSSYSNRRCLITSFLND 514
QY 495 ETYVPKFEAFNAFTPHADICTLSEKROIKKOTAFALVELVKKPKATKEOLKAYMDPFAF 554
Db 515 ETYAPPFSKDFEFHKOQAOQALQTMKQELLINLVKORPELTBEQLAAVADFSL 574
QY 555 VEKCKADKCTCFEAFEGKILVAASQAALGL 585
Db 575 LEKCKAQDQCVCFTEEGPKLISKTRDALGV 605

RESULT 12
Q8JIA7
ID Q8JIA7 PRELIMINARY; PRT; 400 AA.
AC Q8JIA7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE A/B over-sized serum albumin (fragment).
OS Sphenodon punctatus (Hatteria) (Tuatara).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Sphenodontia; Sphenodontidae; Sphenodon.
OX NCBI_TaxID=8508;
RN [1]
RP SEQUENCE FROM N.A.
RA Metcalf V.J., Brennan S.O., George P.M., Chambers G.K.;
RT "Partial mRNA sequence for tuatara A/B serum albumin";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375973; AAM46106.1; -
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 2.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 2.
DR PROSITE; PS00212; ALBUMIN; 2.
FT NON_TER 1
SQ SEQUENCE 400 AA; 45715 MW; 8DE20609657CF753 CRC64;

Query Match 33.7%; Score 1045; DB 13; Length 400;
Best Local Similarity 47.7%; Pred. No. 5.1e-74;
Matches 186; Conservative 80; Mismatches 124; Indels 0; Gaps 0;

QY 196 QELKCSLQKFGRAKAWARLISORPPKAEFAEVSKLVTDLTKVHTCECHGDLLECAD 255
Db 3 EKHSGFLKSGERAFQANKVLQSLQKFPKAPPEELHKVVLVATKLQKCECHGDMIECLD 62
QY 256 DRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPADLPSLAADFVESKDV 315
Db 63 DRVEMAYICSKQAVFSKIDCKCEKPIVDREVCILQADLDEKPADLPSIAGQYIESTEV 122
QY 316 CENYAEAKDVPFGNFIYEYARRHPDYSVVLRLRLAKTYETTLKCCAAADPHECYAKVD 375
Db 123 KKHVEGKDVLAHFVYEYSRHPFESSQMLRTGKYQDTLDCKCKTENPEPCYKAGE 182
QY 376 EFKPLVEEPONLQKONCELFQEGYKFNALVRRTKKVPQVSTFTIVSVRLNGKVG 435
Db 183 ELARIHQESQELKTHCSFTYTSQGDPEQKMLVRYTKKMPQLPAEELIETSKLTGVG 242
QY 436 KCKHPKAKMPCADYLSVLNQLCVLHEKTPVSDRVYKCTESLVNRRCFCSALEVDE 495
Db 243 KCCPLSEDKRLSCSEKLSWLFECROHEASPVNNHVTCCCTDSYSEMRFCFTKLGVDD 302
QY 496 TVVPKEFNAFTPHADICTLSEKROIKKOTAFALVELVKKPKATKEOLKAYMDPFAFV 555
Db 303 SVYPEFCPSFTFLDEQLCTAPEARLKKQLTFLVLKLIQLKPOIEQLKLVITYDNAME 362
QY 556 EKCKADKCTCFEAFEGKILVAASQAALGL 585
Db 363 EKCCQAKNQCFTTEGKLTQEGKALLGV 392

```

RESULT 13

```

Q8YGH6
ID Q8YGH6 PRELIMINARY; PRT; 603 AA.
AC Q8YGH6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Serum albumin precursor (fragment).
GN ALB.
OS Rana shqiperica.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=44326;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bushat; TISSUE=Liver;
RA Uzzell T., Hotz H.;
RT "Albumin cDNA sequence of Rana shqiperica: evolutionary changes in
RT frog albumins.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U04052; AAD09358.1; -
DR HSSP; P02768; 1E7B.
DR InterPro; IPR000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PR00802; SERUMALBUMIN.
DR ProDom; PD002486; Serum_albumin; 1.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 3.
KW SIGNAL.
FT NON_TER 1
FT SIGNAL <1 23 POTENTIAL.
FT CHAIN 24 603 SERUM ALBUMIN.
FT SEQUENCE 603 AA; 69293 MW; 340D3733FA010C99 CRC64;

Query Match 30.8%; Score 955; DB 13; Length 603;
Best Local Similarity 33.6%; Pred. No. 9.9e-67;
Matches 190; Conservative 119; Mismatches 241; Indels 16; Gaps 3;

QY 12 KDIGENFALVLIAPAQYIQCCPFEDHVKLVNVEYTFEAKTCVADSAENCKSLHTFLG 71
Db 37 KAVGKFAVEKLVLMVAQDFEKLSDHLKVOAKIEADVNCNEKHPKAEKCKPAEILYH 96
QY 72 DKLCVTATREYEGMADCCAKOPERNCEFLQKDDNPPLRVLVR-----PEVDVMT 125
Db 97 DIVCKEEDIDQLYPTWTECCGKAERTKCFYEHE-----VRVEEKIPNIESCK 148
QY 126 AFHNEETLKYLYEIAARRHPYAPPELLFTAKRYKAFTPECCOAAADKACLLKDL 185
Db 149 EHKEHFPQAFSYLSNIAKRSKLYPPAVLGFAIQYNEITTECCAAEDKACFGEMPOV 208
QY 186 RDEKASSAKQRLKCSQKFGRAKAWARLISORPPKAEFAEVSKLVTDLTKVHTEC 245
Db 209 KKLNTYLEDKHKQCEVLEKPEPVSQALTIVQVSQFGNAKYDDVEKVTIEAHLNEDC 268
QY 246 CHGDLLECADRADLAKYICENQDSISSKLEKCEKPLEKSHCIAEVENDEMPADLPSL 305
Db 269 CKGDVECMIEEMATEICIAKEKLSKLSDDCAKGVLETPCILALPNEE--PDLPTE 326
QY 306 AADPVSQKVCNAYAEADKDFLGMFLYEYARRHPDYSVVLRLRLAKTYETTLKCCAAAD 365
Db 327 LKEYTEDEHVCNENQDKARKTYLAHTHEDYSRSHQSSQSCILNVSQFGFEMLEKCCASN 386
QY 366 PHECYAKVDKPLVEEPONLQKONCELFQEGYKFNALVRRTKKVPQVSTFTIV 425
Db 387 SAECLEKAPKLEAALKENEETISKONGALEKLGNDFFVIQLLVRYTKMPQVTAQTIVE 446
QY 426 VSRNLGKVSCKCKPIEAKRMPCADYLSVLNQLCVLHEKTPVSDRVYKCTESLVNRR 485
Db 447 LTGEMAKIGYCCGLPDNKKQPCAEKLDILLGEMCEREKKTFFINDNVHCCYDSTANRR 506
QY 486 PCFSALEVDFTYVPKEFNAFTPHADICTLSEKROIKKOTAFALVELVKKPKATKEOLK 545
Db 507 PCFTKGLGPYANPAVPWDESKLHFTDMCKGSADQLKTLKLVLEFKMRKPTCGKEKLT 566

```

Db 344 CTTTTEQRNNYKESFTLRNHPKSLDLEILYKYKLEECQSEHHVQCLHGGEQ 403

QY 532 LVKHKPKATKEQLKAVMDPAAFVEKCKKADDKETCFAEKGKLVAA5QAAAL 583
:||||| | |||||: || ||||| : : || || || : ||||
Db 354 VVKCKPATTHEQLKAVITDFYGVVEKCKCHGENHEACFLAEGPQLVQFTQAAAL 405

Search completed: August 31, 2003, 16:36:34
Job time : 111 secs